

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 18:06:39 ; Search time 76 Seconds
(without alignments)
489.170 Million cell updates/sec

Title: US-10-007-267a-8

Perfect score: 1461

Sequence: 1 MGNHVSLSAERRAHIA.....REKRQRREQLGKIVPEQ 279

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 segs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: A.Geneseq.101002.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
24: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1461	100.0	279	17	AA91312
2	1288	88.2	275	20	AAV22156
3	1036.5	70.9	280	17	AA91315
4	1031.5	70.6	280	18	AAW0580
5	219	15.0	50	20	AAW89328
6	217	14.9	50	20	AAW89329
7	129.5	8.9	423	23	ABB89453
8	129.5	8.9	517	21	AAW41790
9	129.5	8.9	517	22	AAW93202
10	129.5	8.9	517	22	AAW93206

11	129.5	8.9	578	22	AA681257
12	129.5	8.9	579	20	AAV30812
13	129.5	8.9	595	21	AAW12121
14	129.5	8.9	636	23	ABP42016
15	121	8.3	1128	22	ABW62795
16	119.5	8.2	521	21	AAW41900
17	119.5	8.2	622	22	AAW93427
18	119.5	8.2	622	22	AAW88452
19	119.5	8.2	622	22	ABP43477
20	106	7.3	201	18	AAW20604
21	105.5	7.2	168	22	ABG22090
22	99.5	6.8	273	22	ABW60317
23	98.5	6.7	273	20	AAV32543
24	97.5	6.7	273	22	AAW60321
25	96.5	6.6	578	21	AAV97423
26	95.5	6.5	161	23	AAE15943
27	95.5	6.5	389	22	ABW82459
28	94	6.4	739	22	ABW11417
29	93.5	6.4	444	23	AAW76669
30	90	6.2	261	22	ABW03682
31	89.5	6.1	1378	23	ABW57176
32	88.5	6.1	159	21	AAW41784
33	88.5	6.1	500	22	AAU27796
34	88	6.0	651	20	AAV35171
35	87	6.0	279	18	AAW21023
36	86.5	5.9	852	22	ABW08255
37	86.5	5.9	955	14	AAW42088
38	86.5	5.9	955	14	AAW42234
39	86.5	5.9	955	14	AAW42235
40	86.5	5.9	955	14	AAW42236
41	86.5	5.9	955	14	AAW42237
42	86.5	5.9	955	14	AAW42238
43	86.5	5.9	955	14	AAW42239
44	86.5	5.9	955	14	AAW42240
45	86.5	5.9	955	14	AAW42241

ALIGNMENTS

RESULT 1	AA91312
ID	AA91312 standard; Protein; 279 AA.
XX	AC
XX	AA91312;
XX	09-JUL-1996 (first entry)
XX	DE
XX	N. gonorrhoeae glycosyltransferase LgtB.
XX	KW
KW	glycosyltransferase; lipo-oligosaccharide; lgt gene; LOS locus;
KW	vaccine.
XX	OS
XX	Neisseria gonorrhoeae strain F62.
XX	PN
XX	W09610086-A1.
XX	PD
XX	04-APR-1996.
XX	PF
XX	25-SEP-1995; 95WO-US12317.
XX	PR
XX	26-SEP-1994; 9AUS-0312387.
XX	PA
XX	(UYRQ) UNIV ROCKEFELLER.
XX	PI
XX	Gotschlich EC;
XX	WPI; 1996-200924/20.
XX	DR
XX	P-PSDB; AAT14061.
XX	PT
PT	Nucleic acids encoding glycosyl transferase(s) - used in the
PT	diagnosis of infection with Neisseria and for the biosynthesis of
PT	oligo:saccharide(s)

XX Claim 9; Fig 2c; 81pp; English.
 PS
 CC 5 Glycosyltransferases (AAR91311-15) are products of the 1gt locus
 CC (AAR14061) of *Neisseria gonorrhoeae* strain F62. Glycosyltransferase
 CC LgtB (AAR91312) can be obtd. by expression of the lgtB coding
 CC sequence in recombinant host cells. A method for adding gal
 CC betail-4 to GlcNAc or Glc comprises contacting a reaction mixture
 CC contg. activated gal to an acceptor moiety comprising a GlcNAc
 CC or Glc residue in the presence of LgtB. Oligosaccharides can be
 CC produced that, when attached to non-toxic lipids, are useful for
 CC *Neisseria* vaccine prepn. Blood group core oligosaccharides, and
 CC mutants of lacto-N-neotetrase, gangliosides and saccharide
 CC portions of globoglycolipids can also be produced using the enzymes.
 CC
 SQ Sequence 279 AA;
 Query Match 100.0%; Score 1461; DB 17; Length 279;
 Best Local Similarity 100.0%; Pred. No. 4.6e-159; Mismatches 0; Indels 0; Gaps 0;
 Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MONHVISLSAERRAHIAATFGSGRIGPFFDAMPSERLERAMAEIVPGISAHPTLSG 60
 DB 1 MONHVISLSAERRAHIAATFGSGRIGPFFDAMPSERLERAMAEIVPGISAHPTLSG 60
 QY 61 VERACFMSHAYLMEQALDEGVPIYAVFEDDVLGGAQFLAEDTWLOERDPDSAFYVR 120
 DB 61 VERACFMSHAYLMEQALDEGVPIYAVFEDDVLGGAQFLAEDTWLOERDPDSAFYVR 120
 QY 121 LETMFHVLTPSGVADYGGRAFPLESEHGCTAGYIISRKAMRFFLDRFAVLPPERLHP 180
 DB 121 LETMFHVLTPSGVADYGGRAFPLESEHGCTAGYIISRKAMRFFLDRFAVLPPERLHP 180
 QY 181 VDLMEFGNPDREGMPVQOLNPAQELHYAKFHDQNSALGSLIEHDRRLNRKQWRDS 240
 DB 181 VDLMEFGNPDREGMPVQOLNPAQELHYAKFHDQNSALGSLIEHDRRLNRKQWRDS 240
 QY 241 PANTFKHRLIRALTKIGRERKRORRROLGKTIIVPQ 279
 DB 241 PANTFKHRLIRALTKIGRERKRORRROLGKTIIVPQ 279
 RESULT 2
 ID AAY22156 standard; Protein; 275 AA.
 AC AAY22156;
 DT 08-SEP-1999 (first entry)
 DE N. meningitidis Beta-1,4-galactosyltransferase.
 KW Beta-1,4-galactosyltransferase; lgtb; fusion protein; catalytic domain;
 KW glycosyltransferase; accessory enzyme; nucleotide sugar formation;
 KW saccharide donor; oligosaccharide synthesis;
 KW carbohydrate structure development.
 OS *Neisseria meningitidis*.
 PN WO9931224-A2.
 PD 24-JUN-1999.
 PF 15-DEC-1998; 98WO-CA01180.
 PR 14-DEC-1998; 98US-0211691.
 PR 15-DEC-1997; 97US-0069443.
 PA (CANA) NAT RES COUNCIL CANADA.
 PI Gilbert M, Wakarchuk WM, Young NM;
 DR WPI; 1999-395174/33.

DR N-PSDB; AAX84281.
 XX
 PT A new glycosyltransferase fusion protein useful in the enzymatic
 CC synthesis of oligosaccharides
 PS Example 2; Fig 2; 63pp; English.
 XX
 CC This sequence represents the *Neisseria meningitidis* Beta-1,4-
 CC galactosyltransferase (also referred to as lgtb). The invention relates
 CC to a nucleic acid encoding a fusion protein that comprises a
 CC glycosyltransferase catalytic domain and a catalytic domain from an
 CC accessory enzyme that is involved in formation of a nucleotide sugar
 CC which is a saccharide donor for a glycosyltransferase reaction. The
 CC fusion protein is useful in the enzymatic synthesis of oligosaccharides.
 CC The fusion proteins are able to catalyse more than one reaction involved
 CC in the enzymatic synthesis. This is useful for the development of
 CC therapeutic agents that have specific carbohydrate structures.
 CC Carbohydrates are involved in recognition elements on the surface of
 CC cells. The fusion protein can be used for the synthesis of both natural
 CC carbohydrates and synthetic derivatives with novel properties. The fusion
 CC polypeptide allows two glycosyltransferase reactions in a single vessel,
 CC provides improved yields of end products. Additionally, cleanup and
 CC disposal of extra solvents and by-products is reduced. The fusion protein
 CC can also use directly different donor analogues and various acceptors
 CC with a terminal galactose residue.
 XX
 SQ Sequence 275 AA;
 Query Match 88.2%; Score 1288; DB 20; Length 275;
 Best Local Similarity 90.0%; Pred. No. 3.4e-139; Mismatches 17; Indels 4; Gaps 1;
 Matches 251; Conservative 7; Mismatches 17; Indels 4; Gaps 1;
 QY 1 MONHVISLSAERRAHIAATFGSGRIGPFFDAMPSERLERAMAEIVPGISAHPTLSG 60
 DB 1 MONHVISLSAERRAHIAATFGSGRIGPFFDAMPSERLERAMAEIVPGISAHPTLSG 60
 QY 61 VERACFMSHAYLMEQALDEGVPIYAVFEDDVLGGAQFLAEDTWLOERDPDSAFYVR 120
 DB 61 VERACFMSHAYLMEQALDEGVPIYAVFEDDVLGGAQFLAEDTWLOERDPDSAFYVR 120
 QY 121 LETMFHVLTPSGVADYGGRAFPLESEHGCTAGYIISRKAMRFFLDRFAVLPPERLHP 180
 DB 121 LETMFHVLTPSGVADYGGRAFPLESEHGCTAGYIISRKAMRFFLDRFAVLPPERLHP 180
 QY 181 VDLMEFGNPDREGMPVQOLNPAQELHYAKFHDQNSALGSLIEHDRRLNRKQWRDS 240
 DB 181 VDLMEFGNPDREGMPVQOLNPAQELHYAKFHDQNSALGSLIEHDRRLNRKQWRDS 240
 QY 241 PANTFKHRLIRALTKIGRERKRORRROLGKTIIVPQ 279
 DB 241 PANTFKHRLIRALTKIGRERKRORRROLGKTIIVPQ 275
 RESULT 3
 ID AAR91315 standard; Protein; 280 AA.
 AC AAR91315;
 DT 09-JUL-1996 (first entry)
 DE N. gonorrhoeae glycosyltransferase lgtE.
 KW Glycosyltransferase; lipo-oligosaccharide; lgt gene; LOS locus;
 KW vaccine.
 OS *Neisseria gonorrhoeae* strain F62.
 PN WO9610086-A1.
 PD 04-APR-1996.
 PR 25-SEP-1995; 95WO-US12317.


```

XX 26-SEP-1994: 94US-0312387.
PR (UYRQ ) UNIV ROCKEFELLER.
XX
XX Gotschlich EC;
PI
XX WPI; 1996-200924/20.
DR N-PSDB; AAT14061.
XX
PT Nucleic acids encoding glycosyl transferase(s) - used in the
PT diagnosis of infection with Neisseria and for the biosynthesis of
PT oligosaccharide(s)
XX
XX Claim 12; Fig 2f; 81pp; English.
XX
XX 5 Glycosyltransferases (AAR91311-15) are products of the igt locus
XX (AAT14061) of Neisseria gonorrhoeae strain PG2. Glycosyltransferase
XX 1gte (AAR91315) can be obtd. by expression of the 1gte coding
XX sequence in recombinant host cells. A method for adding gal
XX beta1-4 to GlcNAc or Glc comprises contacting a reaction mixture
XX contg. activated gal to an acceptor moiety comprising a GlcNAc or
XX Glc. residue in the presence of 1gte. Oligosaccharides can be produced
XX that, when attached to non-toxic lipids, are useful for Neisseria
XX vaccine prepn. Blood group core oligosaccharides, and mimics of
XX lacto-N-neotetraose, gangliosides and saccharide portions of
XX globoglycolipids can also be produced using the enzymes.
XX
SQ Sequence 280 AA;

Query Match 70.9%; Score 1036.5; DB 17; Length 280;
Best Local Similarity 71.6%; Pred. No. 2.8e-110;
Matches 209; Conservative 24; Mismatches 34; Indels 25; Gaps 3;

QY 1 MGNHVISLASAERRAHIAATFGSRGIPFQFDALMPSERLERAMAEVGLSAHPYLSG 60
DB 1 MGNHVISLASAERRAHIAATFGSRGIPFQFDALMPSERLERAMAEVGLSAHPYLSG 60
QY 61 VERACFMSHAYLWEQALDEGPYIAVEEDVLLGEGAEQFLAEDTWLEERFDKDSATVR 120
DB 61 VERACFMSHAYLWEQALDEGPYIAVEEDVLLGEGAEQFLAEDTWLEERFDKDSATVR 120
QY 121 LETMFHVLSPSGVADYGRAPFLSEHCCTAGYIISRKAMRFLDRRAVLPPELHP 180
DB 121 LETMFHVLSPSGVADYGRAPFLSEHCCTAGYIISRKAMRFLDRRAVLPPELHP 180
QY 181 VDLMEGPNPDREGMPYCOLNPALCAQELHYAKFHONSALGSLIEHDLRLNRKQWRDS 240
DB 181 VDLMEGPNPDREGMPYCOLNPALCAQELHYAKFHONSALGSLIEHDLRLNRKQWRDS 240
QY 241 PANTFKHR-----LIRALTKIGRERERKRROR-----EOLIGKIIVPQ 279
DB 241 PANTFKHR-----LIRALTKIGRERERKRROR-----EOLIGKIIVPQ 279
DB 234 -----RHRRSLKVMFDLKRALGKFGREKKRMRORQAELEKYGRVILFK 280

RESULT 4
AAM06580 standard; protein; 280 AA.
XX
XX AAM06580;
AC
XX
XX 21-MAR-1997 (first entry)
DE Lipo-oligosaccharide gene-encoded protein.
XX
XX Polyglycosyltransferase; N-acetylglucosaminyl transferase;
KM N-acetylglucosaminyl transferase; lipo-oligosaccharide.
XX
XX Neisseria gonorrhoeae ATCC 33094.
OS
XX
XX W09640971-A1.
PN
XX
XX 19-DEC-1996.
PD

```

```

XX 03-JUN-1996: 96WO-DS08323.
XX
XX 07-JUN-1995: 95US-0478140.
PR
XX
XX (NEOS-) NEOSE TECHNOLOGIES INC.
XX
XX Buczala SL, Johnson KF, Roth S;
PI
XX WPI; 1997-052351/05.
DR N-PSDB; AAT49230.
XX
XX Transfer of at least 2 saccharide units using
PT polyglycosyltransferase - isolated from N. gonorrhoeae, catalyses
PT the addition of both GlcNAc and GalNAc di:saccharide(s) units to a
PT single galactose moiety
XX
XX Disclosure; Fig 2F-H; 38pp; English.
XX
XX A lipo-oligosaccharide-encoding gene region (AAT49230) of Neisseria
XX gonorrhoeae ATCC 33084 includes coding sequences for 5 proteins
XX (AAM06576-80), one of which (AAM06576) is a polyglycosyltransferase
XX that catalyzes the addition of GlcNAc and GalNAc disaccharides to
XX a galactose moiety. The function of the other 4 proteins is not
XX stated in the specification.
XX
SQ Sequence 280 AA;

Query Match 70.6%; Score 1031.5; DB 18; Length 280;
Best Local Similarity 71.6%; Pred. No. 1.1e-109;
Matches 209; Conservative 23; Mismatches 35; Indels 25; Gaps 3;

QY 1 MGNHVISLASAERRAHIAATFGSRGIPFQFDALMPSERLERAMAEVGLSAHPYLSG 60
DB 1 MGNHVISLASAERRAHIAATFGSRGIPFQFDALMPSERLERAMAEVGLSAHPYLSG 60
QY 61 VERACFMSHAYLWEQALDEGPYIAVEEDVLLGEGAEQFLAEDTWLEERFDKDSATVR 120
DB 61 VERACFMSHAYLWEQALDEGPYIAVEEDVLLGEGAEQFLAEDTWLEERFDKDSATVR 120
QY 121 LETMFHVLSPSGVADYGRAPFLSEHCCTAGYIISRKAMRFLDRRAVLPPELHP 180
DB 121 LETMFHVLSPSGVADYGRAPFLSEHCCTAGYIISRKAMRFLDRRAVLPPELHP 180
QY 181 VDLMEGPNPDREGMPYCOLNPALCAQELHYAKFHONSALGSLIEHDLRLNRKQWRDS 240
DB 181 VDLMEGPNPDREGMPYCOLNPALCAQELHYAKFHONSALGSLIEHDLRLNRKQWRDS 240
QY 241 PANTFKHR-----LIRALTKIGRERERKRROR-----EOLIGKIIVPQ 279
DB 241 PANTFKHR-----LIRALTKIGRERERKRROR-----EOLIGKIIVPQ 279
DB 234 -----RHRRSLKVMFDLKRALGKFGREKKRMRORQAELEKYGRVILFK 280

RESULT 5
AAM89328 standard; peptide; 50 AA.
XX
XX AAM89328;
AC
XX
XX 26-FEB-1999 (first entry)
DE Neisseria meningitidis igtB C-terminal peptide.
XX
XX Neisseria meningitidis; igtC; igtB; beta-1,4-galactosyltransferase;
KM glycosyltransferase; proteolytic enzyme.
XX
XX Neisseria meningitidis.
OS
XX
XX W09854331-A2.
PN
XX
XX 03-DEC-1998.
PD
XX
XX 26-MAY-1998: 98WO-IB00975.
PF

```

XX 27-MAY-1997; 97US-0047751.
 XX (CANADA) NAT RES COUNCIL CANADA.
 XX
 XX Wakarchuk WW, Young NM;
 XX WPI; 1999-035177/03.
 XX
 XX Expressing high levels of glycosyltransferases - comprises use of
 PT either host cells deficient in proteolytic enzymes or modified
 PT glycosyltransferase genes deleted in a proteolytic recognition site
 PS
 XX Example 1; Fig 8; 61pp; English.
 XX
 CC A method has been developed of expressing a glycosyltransferase in a
 CC host cell. The method comprises introducing into the host cell a nucleic
 CC acid encoding the glycosyltransferase and incubating the host cell under
 CC conditions appropriate for expression of the glycosyltransferase, where
 CC the host cell substantially lacks a protease that cleaves polypeptides
 CC between two consecutive positively charged amino acid residues. The
 CC oligosaccharide structures which are potential therapeutic agents for
 CC use in the manipulation of cell-cell recognition events, particularly
 CC adhesion of bacteria and viruses to mammalian cells and leukocyte-
 CC endothelial cell interaction through selection in inflammation. The
 CC method provides more readily recoverable active glycosyltransferases
 CC than prior art methods involving mammalian glycosyltransferases. The
 CC present sequence represents a C-terminal peptide from Neisseria
 CC meningitidis 19TB from the present invention.
 CC
 SQ Sequence 50 AA;
 Query Match 15.0%; Score 219; DB 20; Length 50;
 Best Local Similarity 85.2%; Pred. No. 2e-17;
 Matches 46; Conservative 0; Mismatches 4; Indels 4; Gaps 1;
 Oy 226 EHDRLNKQWRDPSANTFKRLRLTKIGRERKRRORRDLGKIIVPQ 279
 Db 1 EHDRLNKQWRDPSANTFKRLRLTKIGRERKRRORRDLGKIIVPQ 279
 RESULT 6
 AAR89328
 ID AAR89329 standard; peptide: 50 AA.
 XX
 AC AAW89329;
 XX
 DT 26-FEB-1999 (first entry)
 XX
 DE Neisseria meningitidis 19TB C-terminal peptide.
 XX
 KW Neisseria meningitidis; 19TC; 19TB; beta-1,4-galactosyltransferase;
 KW glycosyltransferase; proteolytic enzyme;
 XX
 OS Neisseria meningitidis.
 XX
 PN WO9854331-A2.
 XX
 PD 03-DEC-1998.
 XX
 PF 26-MAY-1998; 98WO-IB00975.
 XX
 PR 27-MAY-1997; 97US-0047751.
 XX
 PA (CANADA) NAT RES COUNCIL CANADA.
 XX
 PI Wakarchuk WW, Young NM;
 XX WPI; 1999-035177/03.
 XX
 PT Expressing high levels of glycosyltransferases - comprises use of
 PT either host cells deficient in proteolytic enzymes or modified

PT glycosyltransferase genes deleted in a proteolytic recognition site
 XX
 PS Example 1; Fig 8; 61pp; English.
 XX
 CC A method has been developed of expressing a glycosyltransferase in a
 CC host cell. The method comprises introducing into the host cell a nucleic
 CC acid encoding the glycosyltransferase and incubating the host cell under
 CC conditions appropriate for expression of the glycosyltransferase, where
 CC the host cell substantially lacks a protease that cleaves polypeptides
 CC between two consecutive positively charged amino acid residues. The
 CC oligosaccharide structures which are potential therapeutic agents for
 CC use in the manipulation of cell-cell recognition events, particularly
 CC adhesion of bacteria and viruses to mammalian cells and leukocyte-
 CC endothelial cell interaction through selection in inflammation. The
 CC method provides more readily recoverable active glycosyltransferases
 CC than prior art methods involving mammalian glycosyltransferases. The
 CC present sequence represents a C-terminal peptide from Neisseria
 CC meningitidis 19TB from the present invention.
 CC
 SQ Sequence 50 AA;
 Query Match 14.9%; Score 217; DB 20; Length 50;
 Best Local Similarity 83.3%; Pred. No. 3.4e-17;
 Matches 45; Conservative 0; Mismatches 5; Indels 4; Gaps 1;
 Oy 226 EHDRLNKQWRDPSANTFKRLRLTKIGRERKRRORRDLGKIIVPQ 279
 Db 1 EHDRLNKQWRDPSANTFKRLRLTKIGRERKRRORRDLGKIIVPQ 279
 RESULT 7
 ABB89453
 ID ABB89453 standard; Protein: 423 AA.
 XX
 AC ABB89453;
 XX
 DT 24-MAY-2002 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 1829.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; anti-
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 18-MAY-2001; 2001WO-US16450.
 XX
 PR 19-MAY-2000; 2000US-205515P.
 XX
 PA (HUMAN-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX
 DR WPI; 2002-122018/16.
 XX
 DR N-PSDB; ABL89862.
 XX
 PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive, and
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders -
 PS
 PS Claim 11; SEQ ID NO 1829; 2081pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABL89448-ABL90853) and proteins
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating

CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (anti)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 XX
 SQ Sequence 423 AA:
 Query Match 8.9%; Score 129.5; DB 23; Length 423;
 Best Local Similarity 27.2%; Pred. No. 9.3e-06;
 Matches 65; Conservative 30; Mismatches 91; Indels 53; Gaps 12;
 QY 5 VISLSAERRRHIAATGSRGIPFOFDA----LMPSERLERMAELVPGLSAHY--- 57
 DB 155 VISLARPRDRRRRLASLWEMELSGRVYDAVDGMNLSSAIRNLGVDLLPGYO-DPYSGR 213
 QY 58 -LSGVEKACFMSHAYLMEQALDEGVPIAVFEDVDVLLGSGAQFLAEDTWLOERFPDPSA 116
 DB 214 TLTKEVGCGFLSHYSIWEEVARGIARLVYFEDV-----RF--ESN 253
 QY 117 FVVRLETFMHWLTSP-SCVADYGR--AFPLLESEHCCTAGYIIS-----RKAMRF 166
 DB 254 FGRLERLMEDVEAEKLSMDLYLGRKQVNPKEETAVEGLPGLVVAGSYWTLAVLRLLA 313
 QY 167 LDR--FAVLPPERLHYD---LAMEGNPDDR-----EGMPVCOLNALCAQELHYA 212
 DB 314 GARKLLASQPLRRMLPVDLPLMFDPHNEQYKAHFWRPDLVAFSAOPLLAA-PTHYA 371

RESULT 8
 AAB41790
 ID AAB41790 standard; Protein; 517 AA.
 AC AAB41790;
 XX
 DT 08-FEB-2001 (first entry)
 DE Human ORFX ORF1554 polypeptide sequence SEQ ID NO:3108.
 XX
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnerability; antiparasitic; antiparkinsonian; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antineumatic; antithyroid;
 KW antanaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antinflammatory disease; coagulation;
 KW thrombosis; contraceptive.
 XX
 OS Homo sapiens.
 XX
 PN WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US08621.
 XX

PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CUBA-) CUBAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 XX
 DR WPI: 2000-602362/57.
 DR N-PSDB: AAC75999.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 11; Page 2331-2332; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnerability;
 CC antiparasitic; antiparkinsonian; neurotropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antinflammatory; antibacterial; antiviral; antifungal; antineumatic;
 CC antithyroid; and antanaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 SQ Sequence 517 AA:
 Query Match 8.9%; Score 129.5; DB 21; Length 517;
 Best Local Similarity 27.2%; Pred. No. 1.3e-05;
 Matches 65; Conservative 30; Mismatches 91; Indels 53; Gaps 12;
 QY 5 VISLSAERRRHIAATGSRGIPFOFDA----LMPSERLERMAELVPGLSAHY--- 57
 DB 246 VISLARPRDRRRRLASLWEMELSGRVYDAVDGMNLSSAIRNLGVDLLPGYO-DPYSGR 304
 QY 58 -LSGVEKACFMSHAYLMEQALDEGVPIAVFEDVDVLLGSGAQFLAEDTWLOERFPDPSA 116
 DB 305 TLTKEVGCGFLSHYSIWEEVARGIARLVYFEDV-----RF--ESN 344
 QY 117 FVVRLETFMHWLTSP-SCVADYGR--AFPLLESEHCCTAGYIIS-----RKAMRF 166
 DB 345 FGRLERLMEDVEAEKLSMDLYLGRKQVNPKEETAVEGLPGLVVAGSYWTLAVLRLLA 404
 QY 167 LDR--FAVLPPERLHYD---LAMEGNPDDR-----EGMPVCOLNALCAQELHYA 212
 DB 405 GARKLLASQPLRRMLPVDLPLMFDPHNEQYKAHFWRPDLVAFSAOPLLAA-PTHYA 462

RESULT 9
 AAM93202
 ID AAM93202 standard; Protein; 517 AA.
 AC AAM93202;
 XX
 DT 06-NOV-2001 (first entry)
 DE Human polypeptide, SEQ ID NO: 2591.
 XX


```
0Y      167 IDR--FAVLPRLHPVD-----LMENGPDDDR-----EGMPVCOLNALCAOELHYA 212
        | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db    483 GARKLLASQPLRRMLPEYDEFLPMEHQHNEGYKAFHWPRDLVAFSAQLLAA-PTHTYA 540

RESULT 13
ID AAB12121 standard; Protein; 595 AA.
XX AAB12121;
AC AAB12121;
DT 02-FEB-2001 (first entry)
XX
Hydrophobic domain protein from clone HP02962 isolated from KB cells.
DE
XX Human; secreted protein; membrane protein; hydrophobic domain;
KW proliferation control; differentiation induction; material transport;
KM biophylaxis; signal receptor; ion channel; transporter; immunostimulant;
KV immunosuppressant; haematopoiesis regulator; chemotactic; chemokinetic;
KM haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;
KK autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer.
XX Homo sapiens.
OS WO200029448-AZ.
PN 25-MAY-2000.
XX PF 17-NOV-1999; 99WO-JP06412.
XX PR 17-NOV-1998; 98JP-0326255.
PR 22-DEC-1998; 98JP-0364315.
PR 16-MAR-1998; 98JP-0069811.
PK 27-APR-1999; 98QP-0118299.
PR 19-MAY-1999; 99JP-0138169.
XX PA (SAGA ) SAGAMI CHEM RES CENT.
PA (PROT-) PROTEGENE INC.
XX Kato S, Kimura T;
PI WPI; 2000-387753/33.
XX DR N-PDSB; AAA60183, AAA60193.
XX PT Proteins comprising hydrophobic regions, such as secretory and membrane proteins, useful in research and diagnostics and having various activities e.g. immunomodulatory, antinflammatory, chemokinetic, hemostatic, thrombolytic -
PS Claim 1; Page 184-186; 410pp; English.
CC Secretory proteins play important roles in the proliferation control, the
XX differentiation induction, the maternal transport and the biophysix of
CC cells. Membrane proteins have important roles as signal receptors, ion channels and transporters. The present sequence is a human protein which has at least one hydrophobic domain. This protein may be a secretory or a membrane protein. The present protein may have cytokine and cell proliferation/differentiation activity, immune stimulating or suppressing activity, haematopoiesis activity, tissue growth activity, CC activn/inhibn activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, anti-inflammatory activity and tumour inhibition activity. The present protein could therefore be used for treatment of autoimmune disease, Alzheimer's disease, Parkinson's disease, and cancer.
CC XX
SQ Sequence 595 AA;
```

Db	324	VISLARPRDRRRRLMASLMEKISGRVAVDAVGMNLNSAIRNLGVLLPGYO-DPYSGR	382
Qy	58	-LSGKACQFMSHAVIMWQALDEGVPTIATVEDVYLLGEGAEQPLADTQLQERFPDSDA	116
Db	383	TLTKEVCFSLSHYSIMVEEYVARLARLVAFEDVY-----RF--ESN	422
Qy	117	FVVRLETFEMHVLNSP--SGVADYGR--APFLSEHQTAGYIIS-----RKMAREF	166
Db	423	FRGRLERMEDVEAEKLSWDLITGRKQVNFKEKTAVEGLVGLVAVSYTWTLATVLRLA	482
Qy	167	LDRLFAVLPPERLHPVD---LMMFGNPDDR-----EGMPYCOQNPALCAQELHYA	212
Db	463	GARKLLASQPLRRMLPVDEFPLMFEDQHPNQYKAHWPMBDIVAFSAQPLLA-PTHYA	540
RESULT 14			
ABP42016	ID	ABP42016 standard; Protein; 636 AA.	
XX	AC	ABP42016;	
XX	DT	22-AUG-2002 (first entry)	
XX	DE	Human ovarian antigen HEKHE51, SEQ ID NO:3148.	
XX	XX	Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;	
XX	XX	ovarian cancer; breast cancer; tumour; reproductive system disorder;	
XX	XX	infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;	
XX	XX	PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;	
XX	XX	inflammatory condition; immune disorder; blood disorder;	
XX	XX	cardiovascular disorder; respiratory disorder; neurological disorder;	
XX	XX	gastrointestinal disorder; urinary system disorder; drug screening;	
XX	XX	gene therapy; chromosome mapping; forensic analysis;	
XX	XX	antibody preparation; cytostatic; immunomodulatory; neuroprotective;	
XX	XX	antiinflammatory; gynaecological; reproductive.	
OS		Homo sapiens.	
PN		WO200200677-A1.	
PD		03-JAN-2002.	
PF		07-JUN-2001; 2001WO-US18569.	
PR		07-JUN-2000; 2000US-209467P.	
PA		(HUMA-) HUMAN GENOME SCI INC.	
PI		Birse CE, Rosen CA;	
DR		WPI; 2002-147878/19.	
DR		N-PSDB; ABQ55093.	
PT		Isolated nucleic acid molecules encoding novel ovarian polypeptides,	
PT		useful in the prevention, treatment and diagnosis of cancer (e.g.	
PT		ovarian cancer), immune disorders, cardiovascular disorders and	
PT		neurological diseases -	
PS		Claim 11; SEQ ID NO 3148; 2922pp; English.	
CC		The invention relates to 2175 novel human ovarian antigens (ABP41054-	
CC		ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also	
CC		encompasses polypeptides 90% identical and polynucleotides 95% identical	
CC		to the sequences of the invention. The invention additionally relates to	
CC		recombinant vectors and host cells comprising human ovarian antigen	
CC		polynucleotides, antibodies against human ovarian antigens, and the use	
CC		of ovarian antigen polynucleotides and polypeptides in diagnosing,	
CC		treating, prognosing or preventing various ovary and/or breast-related	
CC		disorders. Such conditions include ovarian cancer and breast cancer, and	
CC		metastatic tumours of ovarian or breast origin, reproductive system	
CC		disorders (e.g., infertility, disorders of pregnancy, anovulation,	
CC		polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine	
CC		disorders, infections (e.g., chlamydia, HIV, toxoplasmosis), and toxic	

CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents a human ovarian antigen of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC
SQ Sequence 636 AA;

Query Match 8.9%; Score 129.5; DB 23; Length 636;
Best Local Similarity 27.2%; Pred. No. 1.7e-05;
Matches 65; Conservative 30; Mismatches 91; Indels 53; Gaps 12;

OY 5 VISLASAERRAHIAATFGSRGIPQFPDA---LMPSERLERAMMELVPGLSAHPY--- 57
DB 365 VISLARPPDRERKRLASLWMEISGRVVDVVGWMLNSAIRNLGYDLFGYQ-DPYSGR 423
OY 58 -LSGVERACFMSHAVLMEQALDEGVPIYAVFEDVLLGEGAEQFLAEDTWLOERFPDPSA 116
DB 424 TLTKGEVGCFLSHYSIMEEVAVARGIARVLVEFDV-----RFE-ESN 463
OY 117 FVRLTETMEHVLITSP-SCVADYGR--APPLISEHCHTAYTIS-----RKAMRFP 166
DB 464 FRRLERLMEVDVAEKLSMDLIYIGRKQVNEKETANEGPLGLVAVGYSYWTLAYALRLA 523
OY 167 LDR--FAVLPPERLHPVD---LMFGNPPDR-----EGMPVCQINPALCAQELHYA 212
DB 524 GARKILASQPLRLKMLPVDLPIFMFDQHPNEQYKAHFWPRDLVAFSAQPLLA-PTHYA 581

RESULT 15
ABB62795
ID ABB62795 standard; protein; 1128 AA.
XX
AC ABB62795;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 15177.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN MO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001MO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PMD, Myers EM;
XX
DR WPI; 2001-656860/75.
XX
DR N-PDB; ABL06898.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions -
XX
PS Disclosure; SEQ ID NO 15177; 21pp + Sequence Listing; English.
XX
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC
SQ Sequence 1128 AA;

Query Match 8.3%; Score 121; DB 22; Length 1128;
Best Local Similarity 24.7%; Pred. No. 0.00038;
Matches 53; Conservative 28; Mismatches 78; Indels 56; Gaps 9;

OY 5 VISLASAERRAHIAATFGSRGIPQFPDA---LMPSERLERAMMELVPGLSAHPY 57
DB 853 MINLKRPRERKMERLDEIGTBAEHFPAVDCKESTERLMEKGVRLFGYEDPYHRA 912
OY 58 -LSGVERACFMSHAVLMEQALDEGVPIYAVFEDVLLGEGAEQFLAEDTWLOERFPDPSA 117
DB 913 MTWGEIGCFLSHYNIMWMMWRKQLEVLLEDI-----RFE----- 950
OY 118 VVRLTETMEHVLITSPSGVADYG---GRAFPLISE-----HCG---TAGYIIS 159
DB 951 --YFRONAVRILNQARNAAQYDLYFGKRRLKESEPAVENADNLVHAGSYWTLYGVIS 1008
OY 160 RKAMRFPDRFAVLPPERLHPVD---LMFGNDP 190
DB 1009 ---LQCALKILAAKPLDKLIPVDLPLMDRHPN 1040

Search completed: May 27, 2003, 18:15:43
Job time : 78 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 18:14:10 ; Search time 29 Seconds
(without alignments)
283.069 Million cell updates/sec

Title: US-10-007-267a-8

Perfect score: 1461

Sequence: 1 MONHVISLASAARRAHIAA.....REKRORRQLGIKTIIVPQ 279

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCFUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfiles1.pep:*

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1461	100.0	279	1	US-08-312-387B-8
2	1461	100.0	279	1	US-08-683-426-8
3	1461	100.0	279	1	US-08-683-458-8
4	1461	100.0	279	2	US-08-878-360-8
5	1461	100.0	279	4	US-09-333-412-8
6	1036.5	70.9	280	1	US-08-312-387B-6
7	1036.5	70.9	280	1	US-08-683-426-6
8	1036.5	70.9	280	1	US-08-683-458-6
9	1036.5	70.9	280	2	US-08-878-360-6
10	1036.5	70.9	280	3	US-08-478-140B-6
11	1036.5	70.9	280	4	US-09-333-412-6
12	1036.5	70.9	280	4	US-09-338-943-6
13	82	5.6	304	2	US-08-701-191A-27
14	81.5	5.6	337	4	US-09-134-001C-3799
15	81	5.5	554	4	US-09-564-805-236
16	80.5	5.5	540	3	US-08-964-268-6
17	80	5.5	285	4	US-09-186-276B-54
18	80	5.5	285	4	US-08-842-445-54
19	80	5.5	285	4	US-09-186-188B-54
20	79.5	5.4	969	1	US-08-365-689-3
21	79.5	5.4	969	1	US-07-747-781-3
22	79.5	5.4	969	1	US-08-145-138A-3
23	79.5	5.4	969	5	PCT-US92-06888-3
24	79.5	5.4	969	5	PCT-US93-03027-4
25	78.5	5.4	316	1	US-08-278-089A-16
26	78.5	5.4	316	2	US-08-838-957A-15
27	78.5	5.4	501	1	US-07-687-466B-2

28	78.5	5.4	501	1	US-08-434-702-2	Sequence 2, Appl1
29	78.5	5.4	501	1	US-08-271-883-2	Sequence 2, Appl1
30	78.5	5.4	501	6	5168064-4	Patent No. 516806
31	75	5.1	628	3	US-08-776-271-2	Sequence 2, Appl1
32	75	5.1	628	4	US-09-215-035-2	Sequence 2, Appl1
33	74.5	5.1	340	1	US-08-606-888A-5	Sequence 5, Appl1
34	73.5	5.0	269	2	US-07-857-224B-79	Sequence 79, Appl1
35	73	5.0	261	1	US-08-245-688-6	Sequence 3, Appl1
36	73	5.0	586	3	US-08-964-268-3	Sequence 3, Appl1
37	73	5.0	738	4	US-08-989-385-1	Sequence 2, Appl1
38	72	4.9	261	1	US-08-245-688-2	Sequence 12, Appl1
39	72	4.9	261	1	US-08-245-688-12	Sequence 12, Appl1
40	71.5	4.9	124	2	US-08-743-200-12	Sequence 12, Appl1
41	71.5	4.9	377	2	US-08-969-106-6	Sequence 6, Appl1
42	71.5	4.9	377	4	US-09-054-492B-1	Sequence 1, Appl1
43	70.5	4.8	256	4	US-09-355-166-2	Sequence 2, Appl1
44	70.5	4.8	329	4	US-09-071-035-368	Sequence 368, App
45	70.5	4.8	357	4	US-09-071-035-366	Sequence 366, App

ALIGNMENTS

RESULT 1
US-08-312-387B-8
Sequence 8, Application US/08312387B
Patent No. 5545553
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESS: Klauder & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/312,387B
FILING DATE: July 7, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 279 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-312-387B-8
Query Match 100.0%; Score 1461; DB 1; Length 279;
Best local Similarity 100.0%; Pred No. 4, 1e-160;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MONHVISLASAARRAHIAATFGSRGIPFOFALPMPSERLERAMETVGLSAHPYLSG 60
Db 1 MONHVISLASAARRAHIAATFGSRGIPFOFALPMPSERLERAMETVGLSAHPYLSG 60
QY 61 VERACPFMSHAVLMEQALDEGVPIYIAVFEDVYLLGEGAEQFLAEDTWLQERFPDPSAFVR 120

Db 61 VERACFMSHAUWMEQALDEGVYIAVFEDDVLGEGAOFLAEDTWLOERDPPSAFYVR 120
QY 121 LETMFHVLTPSPGVADYGGRAFPLESEHCCTAGYIISRKAMREFLDRFAVLPPERLHP 180
Db 121 LETMFHVLTPSPGVADYGGRAFPLESEHCCTAGYIISRKAMREFLDRFAVLPPERLHP 180
QY 181 VDLAMFGNPDDEGMPVQOLPALCAOELHYAKFHDONSALGSLIEHRRRLNRKQOWMDS 240
Db 181 VDLAMFGNPDDEGMPVQOLPALCAOELHYAKFHDONSALGSLIEHRRRLNRKQOWMDS 240
QY 241 PANTFKHRLIRALTIGRERERKRRORREOLIGKIIVPQ 279
Db 241 PANTFKHRLIRALTIGRERERKRRORREOLIGKIIVPQ 279

RESULT 2

US-08-683-426-8
Sequence 8, Application US/08683426
Patent No. 5705367
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauder & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,426
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/312,387
FILING DATE: September 26, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 279 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-683-426-8

Query Match 100.0%; Score 1461; DB 1; Length 279;
Best Local Similarity 100.0%; Pred. No. 4,1e-160;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MONHVISLASAERRAHIAATFGSRGIPFOFDALMPSERLERAMAEIVPGLSAHPYLSG 60
Db 1 MONHVISLASAERRAHIAATFGSRGIPFOFDALMPSERLERAMAEIVPGLSAHPYLSG 60
QY 61 VERACFMSHAUWMEQALDEGVYIAVFEDDVLGEGAOFLAEDTWLOERDPPSAFYVR 120
Db 61 VERACFMSHAUWMEQALDEGVYIAVFEDDVLGEGAOFLAEDTWLOERDPPSAFYVR 120

QY 121 LETMFHVLTPSPGVADYGGRAFPLESEHCCTAGYIISRKAMREFLDRFAVLPPERLHP 180
Db 121 LETMFHVLTPSPGVADYGGRAFPLESEHCCTAGYIISRKAMREFLDRFAVLPPERLHP 180
QY 181 VDLAMFGNPDDEGMPVQOLPALCAOELHYAKFHDONSALGSLIEHRRRLNRKQOWMDS 240
Db 181 VDLAMFGNPDDEGMPVQOLPALCAOELHYAKFHDONSALGSLIEHRRRLNRKQOWMDS 240
QY 241 PANTFKHRLIRALTIGRERERKRRORREOLIGKIIVPQ 279
Db 241 PANTFKHRLIRALTIGRERERKRRORREOLIGKIIVPQ 279

RESULT 3

US-08-683-458-8
Sequence 8, Application US/08683458
Patent No. 5798233
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauder & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,458
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/312,387
FILING DATE: September 26, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 279 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-683-458-8

Query Match 100.0%; Score 1461; DB 1; Length 279;
Best Local Similarity 100.0%; Pred. No. 4,1e-160;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MONHVISLASAERRAHIAATFGSRGIPFOFDALMPSERLERAMAEIVPGLSAHPYLSG 60
Db 1 MONHVISLASAERRAHIAATFGSRGIPFOFDALMPSERLERAMAEIVPGLSAHPYLSG 60
QY 61 VERACFMSHAUWMEQALDEGVYIAVFEDDVLGEGAOFLAEDTWLOERDPPSAFYVR 120
Db 61 VERACFMSHAUWMEQALDEGVYIAVFEDDVLGEGAOFLAEDTWLOERDPPSAFYVR 120
QY 121 LETMFHVLTPSPGVADYGGRAFPLESEHCCTAGYIISRKAMREFLDRFAVLPPERLHP 180
Db 121 LETMFHVLTPSPGVADYGGRAFPLESEHCCTAGYIISRKAMREFLDRFAVLPPERLHP 180

QY 181 VDLAMFGNPDREGMPVQCLNPALCAOELHYAFHNDONSALGSLIEHDRLNKQQRDS 240
| | | | |
DB 181 VDLAMFGNPDREGMPVQCLNPALCAOELHYAFHNDONSALGSLIEHDRLNKQQRDS 240
QY 241 PANTFKHRLIRALTIGRERERKRRORREOLIGKIIVFQ 279
| | | | |
DB 241 PANTFKHRLIRALTIGRERERKRRORREOLIGKIIVFQ 279

RESULT 4

US-08-878-360-8
; Sequence 8, Application US/08878360
; Patent No. 5945322
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,360
; FILING DATE: 18-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/683,426
; FILING DATE:
; APPLICATION NUMBER: 08/312,387
; FILING DATE: September 26, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-0958
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-878-360-8

Query Match 100.0%; Score 1461; DB 2; Length 279;
Best Local Similarity 100.0%; Pred. No. 4.1e-160;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MONHVSIISAARRAHIAATGSRGIPPOFPALMPSERLERAMAEIVPGLSAHPYLSG 60
| | | | |
DB 1 MONHVSIISAARRAHIAATGSRGIPPOFPALMPSERLERAMAEIVPGLSAHPYLSG 60
QY 61 VERACFMSHAVLMEQALDEGVPIYAFVEDVLLGEGAEQFLADDTWLOERFDDPSAFVVR 120
| | | | |
DB 61 VERACFMSHAVLMEQALDEGVPIYAFVEDVLLGEGAEQFLADDTWLOERFDDPSAFVVR 120
QY 121 LETMFHVLTPSPGVADYGGRAFPLLSEHCGTAGYIISRKARFFLDRAVALPPERLHP 180
| | | | |
DB 121 LETMFHVLTPSPGVADYGGRAFPLLSEHCGTAGYIISRKARFFLDRAVALPPERLHP 180

QY 181 VDLAMFGNPDREGMPVQCLNPALCAOELHYAFHNDONSALGSLIEHDRLNKQQRDS 240
| | | | |
DB 181 VDLAMFGNPDREGMPVQCLNPALCAOELHYAFHNDONSALGSLIEHDRLNKQQRDS 240
QY 241 PANTFKHRLIRALTIGRERERKRRORREOLIGKIIVFQ 279
| | | | |
DB 241 PANTFKHRLIRALTIGRERERKRRORREOLIGKIIVFQ 279

RESULT 5

US-09-333-412-8
; Sequence 8, Application US/09333412
; Patent No. 6342382
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/333,412
; FILING DATE: 15-JUN-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/312,387
; FILING DATE: July 7, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-333-412-8

Query Match 100.0%; Score 1461; DB 4; Length 279;
Best Local Similarity 100.0%; Pred. No. 4.1e-160;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MONHVSIISAARRAHIAATGSRGIPPOFPALMPSERLERAMAEIVPGLSAHPYLSG 60
| | | | |
DB 1 MONHVSIISAARRAHIAATGSRGIPPOFPALMPSERLERAMAEIVPGLSAHPYLSG 60
QY 61 VERACFMSHAVLMEQALDEGVPIYAFVEDVLLGEGAEQFLADDTWLOERFDDPSAFVVR 120
| | | | |
DB 61 VERACFMSHAVLMEQALDEGVPIYAFVEDVLLGEGAEQFLADDTWLOERFDDPSAFVVR 120
QY 121 LETMFHVLTPSPGVADYGGRAFPLLSEHCGTAGYIISRKARFFLDRAVALPPERLHP 180
| | | | |
DB 121 LETMFHVLTPSPGVADYGGRAFPLLSEHCGTAGYIISRKARFFLDRAVALPPERLHP 180
QY 181 VDLAMFGNPDREGMPVQCLNPALCAOELHYAFHNDONSALGSLIEHDRLNKQQRDS 240
| | | | |
DB 181 VDLAMFGNPDREGMPVQCLNPALCAOELHYAFHNDONSALGSLIEHDRLNKQQRDS 240

OY 241 PANTFKHRLRALFKIGREKRRQRRLQIGKIIVPQ 279
Db 241 PANTFKHRLRALFKIGREKRRQRRLQIGKIIVPQ 279

RESULT 6

US-08-312-387B-6
; Sequence 6, Application US/08312387B
; Patent No. 5545553

GENERAL INFORMATION:

APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
NUMBER OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/312,387B
FILING DATE: July 7, 1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-0955
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-312-387B-6

Query Match 70.9%; Score 1036.5; DB 1; Length 280;
Best Local Similarity 71.6%; Pred. No. 3.4e-111;

Matches 209; Conservative 24; Mismatches 34; Indels 25; Gaps 3;

OY 1 MONHVISLSAASERRAHIAATFGSGIRPQFDALMPSERLERAMAEIVPGLSAHPYLSG 60
Db 1 MONHVISLSAASERRAHIAATFGSGIRPQFDALMPSERLERAMAEIVPGLSAHPYLSG 60
OY 61 VEKACFMSHAIVLMQALDEGIPYIAVFEDDVILGSGAQFLAEDWTLOERDPPDSAFYVR 120
Db 61 VEKACFMSHAIVLMQALDEGIPYIAVFEDDVILGSGAQFLAEDWTLOERDPPDSAFYVR 120
OY 121 LETFMHVLTPSPGVADYGRAPFLSEHGCTAGYIISRKAMFFLDRFAVLPERLHP 180
Db 121 LETFMHVLTPSPGVADYGRAPFLSEHGCTAGYIISRKAMFFLDRFAVLPERLHP 180
OY 181 VDLMEFGNPDREGMPVQQLPALCAQELAHAKFDONSALGSLIEHDRIRNRKOQWRDS 240
Db 181 VDLMEFGNPDREGMPVQQLPALCAQELAHAKFDONSALGSLIEHDRIRNRKOQWRDS 240
OY 241 PANTFKHR-----LIRALTKIGREKRRQR-----EOLIGKIIVPQ 279
Db 241 PANTFKHR-----LIRALTKIGREKRRQR-----EOLIGKIIVPQ 279

RESULT 7

US-08-683-426-6
; Sequence 6, Application US/08683426
; Patent No. 5705367

GENERAL INFORMATION:

APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
NUMBER OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,426
FILING DATE:
CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/312,387
FILING DATE: September 26, 1994

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-0955B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-683-426-6

Query Match 70.9%; Score 1036.5; DB 1; Length 280;
Best Local Similarity 71.6%; Pred. No. 3.4e-111;
Matches 209; Conservative 24; Mismatches 34; Indels 25; Gaps 3;

OY 1 MONHVISLSAASERRAHIAATFGSGIRPQFDALMPSERLERAMAEIVPGLSAHPYLSG 60
Db 1 MONHVISLSAASERRAHIAATFGSGIRPQFDALMPSERLERAMAEIVPGLSAHPYLSG 60
OY 61 VEKACFMSHAIVLMQALDEGIPYIAVFEDDVILGSGAQFLAEDWTLOERDPPDSAFYVR 120
Db 61 VEKACFMSHAIVLMQALDEGIPYIAVFEDDVILGSGAQFLAEDWTLOERDPPDSAFYVR 120
OY 121 LETFMHVLTPSPGVADYGRAPFLSEHGCTAGYIISRKAMFFLDRFAVLPERLHP 180
Db 121 LETFMHVLTPSPGVADYGRAPFLSEHGCTAGYIISRKAMFFLDRFAVLPERLHP 180
OY 181 VDLMEFGNPDREGMPVQQLPALCAQELAHAKFDONSALGSLIEHDRIRNRKOQWRDS 240
Db 181 VDLMEFGNPDREGMPVQQLPALCAQELAHAKFDONSALGSLIEHDRIRNRKOQWRDS 240
OY 241 PANTFKHR-----LIRALTKIGREKRRQR-----EOLIGKIIVPQ 279
Db 241 PANTFKHR-----LIRALTKIGREKRRQR-----EOLIGKIIVPQ 279

RESULT 8

US-08-683-458-6
; Sequence 6, Application US/08683458
; Patent No. 5798233

GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES: 12
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,458
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/312,387
FILING DATE: September 26, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-683-458-6

Query Match 70.9%; Score 1036.5; DB 1; Length 280;
Best Local Similarity 71.6%; Pred. No. 3.4e-111;
Matches 209; Conservative 24; Mismatches 34; Indels 25; Gaps 3;

1 MONHVISLASAERRRAIATFGSRGIPFOFDALMPSERLERAMAEIVGLSAHPYLSG 60
1 MONHVISLASAERRRAIATFGSRGIPFOFDALMPSERLERAMAEIVGLSAHPYLSG 60
61 VERACFMSHAVLMEQALDEGVPIYAVEEDVLLGEGAEQFLAEDTWLEERPDKDSAFIVR 120
61 VERACFMSHAVLMEQALDEGVPIYAVEEDVLLGEGAEQFLAEDTWLEERPDKDSAFIVR 120
121 LETMFHNVLTSPSGVADYGRAPFLPSEHCCTAGYIISRKAMFFLDRAVLPPELHP 180
121 LETMFHNVLTSPSGVADYGRAPFLPSEHCCTAGYIISRKAMFFLDRAVLPPELHP 180
121 LETMFHNVLTSPSGVADYGRAPFLPSEHCCTAGYIISRKAMFFLDRAVLPPELHP 180
121 LETMFHNVLTSPSGVADYGRAPFLPSEHCCTAGYIISRKAMFFLDRAVLPPELHP 180
181 VDLAMFNPDRGMPVPCOLNPAALCAQELHYAKFHONSAIGSLIEHRRDLNRKQGRDS 240
181 VDLAMFNPDRGMPVPCOLNPAALCAQELHYAKFHONSAIGSLIEHRRDLNRKQGRDS 240
181 VDLAMFNPDRGMPVPCOLNPAALCAQELHYAKFHONSAIGSLIEHRRDLNRKQGRDS 240
181 VDLAMFNPDRGMPVPCOLNPAALCAQELHYAKFHONSAIGSLIEHRRDLNRKQGRDS 240
241 PANTFKR-----LIRALTKIGRERERKRR-----EOLIGKTIIVPEQ 279
241 PANTFKR-----LIRALTKIGRERERKRR-----EOLIGKTIIVPEQ 279
234 -----RRRRLSKVMEFLKRALGKFGREKKRMERQDALEKYGRAVILFK 280
234 -----RRRRLSKVMEFLKRALGKFGREKKRMERQDALEKYGRAVILFK 280

RESULT 9
US-08-878-360-6
Sequence 6, Application US/08878360
Patent No. 5945322
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF

TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES: 12
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,360
FILING DATE: 18-JUN-1997
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/683,426
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/312,387
FILING DATE: September 26, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-878-360-6

Query Match 70.9%; Score 1036.5; DB 2; Length 280;
Best Local Similarity 71.6%; Pred. No. 3.4e-111;
Matches 209; Conservative 24; Mismatches 34; Indels 25; Gaps 3;

1 MONHVISLASAERRRAIATFGSRGIPFOFDALMPSERLERAMAEIVGLSAHPYLSG 60
1 MONHVISLASAERRRAIATFGSRGIPFOFDALMPSERLERAMAEIVGLSAHPYLSG 60
61 VERACFMSHAVLMEQALDEGVPIYAVEEDVLLGEGAEQFLAEDTWLEERPDKDSAFIVR 120
61 VERACFMSHAVLMEQALDEGVPIYAVEEDVLLGEGAEQFLAEDTWLEERPDKDSAFIVR 120
121 LETMFHNVLTSPSGVADYGRAPFLPSEHCCTAGYIISRKAMFFLDRAVLPPELHP 180
121 LETMFHNVLTSPSGVADYGRAPFLPSEHCCTAGYIISRKAMFFLDRAVLPPELHP 180
121 LETMFHNVLTSPSGVADYGRAPFLPSEHCCTAGYIISRKAMFFLDRAVLPPELHP 180
121 LETMFHNVLTSPSGVADYGRAPFLPSEHCCTAGYIISRKAMFFLDRAVLPPELHP 180
181 VDLAMFNPDRGMPVPCOLNPAALCAQELHYAKFHONSAIGSLIEHRRDLNRKQGRDS 240
181 VDLAMFNPDRGMPVPCOLNPAALCAQELHYAKFHONSAIGSLIEHRRDLNRKQGRDS 240
181 VDLAMFNPDRGMPVPCOLNPAALCAQELHYAKFHONSAIGSLIEHRRDLNRKQGRDS 240
181 VDLAMFNPDRGMPVPCOLNPAALCAQELHYAKFHONSAIGSLIEHRRDLNRKQGRDS 240
241 PANTFKR-----LIRALTKIGRERERKRR-----EOLIGKTIIVPEQ 279
241 PANTFKR-----LIRALTKIGRERERKRR-----EOLIGKTIIVPEQ 279
234 -----RRRRLSKVMEFLKRALGKFGREKKRMERQDALEKYGRAVILFK 280
234 -----RRRRLSKVMEFLKRALGKFGREKKRMERQDALEKYGRAVILFK 280

RESULT 10
US-08-478-140B-6
Sequence 6, Application US/08478140B
Patent No. 6127153
GENERAL INFORMATION:
APPLICANT: JOHNSON, KARL F.
APPLICANT: ROTH, STEPHEN
APPLICANT: BUCZALA, STEPHANIE L.

```

; TITLE OF INVENTION: METHOD OF TRANSFERRING AT LEAST TWO
; TITLE OF INVENTION: SACCHARIDE UNITS WITH A POLYGLYCOSYLTRANSFERASE, A
; TITLE OF INVENTION: POLYGLYCOSYLTRANSFERASE AND GENE ENCODING A
; TITLE OF INVENTION: POLYGLYCOSYLTRANSFERASE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,140B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Laura A. Coruzzi
; REGISTRATION NUMBER: 30,742
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELE: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 280 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-478-140B-6

```

```

Query Match 70.9%; Score 1036.5; DB 3; Length 280;
Best Local Similarity 71.6%; Pred. No. 3.4e-111;
Matches 209; Conservative 24; Mismatches 34; Indels 25; Gaps 3;

```

```

QY 1 MONHYSLSAASERRAHNAATPGSRGIPQFPDAMPSERLERAMAEVPGSAHPYLSG 60
DB- 1 MONHYSLSAASERRAHNAATPGSRGIPQFPDAMPSERLERAMAEVPGSAHPYLSG 60
QY 61 VEKACFMSHAVLMEQALDEGVPIYAFVEDDVLGGAQOFLAEDTWLOERFPDPSAFYVR 120
DB 61 VEKACFMSHAVLMEQALDEGVPIYAFVEDDVLGGAQOFLAEDTWLOERFPDPSAFYVR 120
QY 121 LETMFAKVIYRDKVLTENKSFPLESEHCGTAGIISRKAMRFLDRFAVLPPERLHP 180
DB 121 LETMFAKVIYRDKVLTENKSFPLESEHCGTAGIISRKAMRFLDRFAVLPPERLHP 180
QY 181 VDLMEFGNPDREGMPVQOLNAPALCAOELHYAKFHDONSALGSLTEHDRRLRKQOWMS 240
DB 181 VDLMEFGNPDREGMPVQOLNAPALCAOELHYAKFHDONSALGSLTEHDRRLRKQOWMS 240
QY 241 PANTFKHR-----LIRALFKIGREKERKRR-----EOLICKIIVPQ 279
DB 241 PANTFKHR-----LIRALFKIGREKERKRR-----EOLICKIIVPQ 279
QY 234 -----RHRRSLKVMEDLKRALGKFGREKKRMEROROAELKRYGRVILFK 280
DB 234 -----RHRRSLKVMEDLKRALGKFGREKKRMEROROAELKRYGRVILFK 280

```

```

RESULT 11
US-09-333-412-6
; Sequence 6, Application US/09333412
; Patent No. 6342382
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson

```

```

; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/333,412
; FILING DATE: 15-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/312,387
; FILING DATE: July 7, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELE: 133521
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 280 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
; US-09-333-412-6

```

```

Query Match 70.9%; Score 1036.5; DB 4; Length 280;
Best Local Similarity 71.6%; Pred. No. 3.4e-111;
Matches 209; Conservative 24; Mismatches 34; Indels 25; Gaps 3;

```

```

QY 1 MONHYSLSAASERRAHNAATPGSRGIPQFPDAMPSERLERAMAEVPGSAHPYLSG 60
DB 1 MONHYSLSAASERRAHNAATPGSRGIPQFPDAMPSERLERAMAEVPGSAHPYLSG 60
QY 61 VEKACFMSHAVLMEQALDEGVPIYAFVEDDVLGGAQOFLAEDTWLOERFPDPSAFYVR 120
DB 61 VEKACFMSHAVLMEQALDEGVPIYAFVEDDVLGGAQOFLAEDTWLOERFPDPSAFYVR 120
QY 121 LETMFAKVIYRDKVLTENKSFPLESEHCGTAGIISRKAMRFLDRFAVLPPERLHP 180
DB 121 LETMFAKVIYRDKVLTENKSFPLESEHCGTAGIISRKAMRFLDRFAVLPPERLHP 180
QY 181 VDLMEFGNPDREGMPVQOLNAPALCAOELHYAKFHDONSALGSLTEHDRRLRKQOWMS 240
DB 181 VDLMEFGNPDREGMPVQOLNAPALCAOELHYAKFHDONSALGSLTEHDRRLRKQOWMS 240
QY 241 PANTFKHR-----LIRALFKIGREKERKRR-----EOLICKIIVPQ 279
DB 241 PANTFKHR-----LIRALFKIGREKERKRR-----EOLICKIIVPQ 279
QY 234 -----RHRRSLKVMEDLKRALGKFGREKKRMEROROAELKRYGRVILFK 280
DB 234 -----RHRRSLKVMEDLKRALGKFGREKKRMEROROAELKRYGRVILFK 280

```

```

RESULT 12
US-09-338-943-6
; Sequence 6, Application US/09338943
; Patent No. 6379933
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, KARL F.
; APPLICANT: BOTH, STEPHEN
; APPLICANT: BOCEZLA, STEPHANIE L.
; TITLE OF INVENTION: METHOD OF TRANSFERRING AT LEAST TWO
; TITLE OF INVENTION: SACCHARIDE UNITS WITH A POLYGLYCOSYLTRANSFERASE, A
; TITLE OF INVENTION: POLYGLYCOSYLTRANSFERASE AND GENE ENCODING A
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:

```

```

ADDRESS: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/338,943
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/478,140
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Laura A. Coruzzi
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7188-017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-338-943-6

Query Match          70.9%; Score 1036.5; DB 4; Length 280;
Best Local Similarity 71.6%; Pred. No. 3.4e-111;
Matches 209; Conservative 24; Mismatches 34; Indels 25; Gaps 3;

QY 1 MONHVISLSAAERRAHIAATFGSGRIGPQFFDALMPSERLERMAELVPGLSAHPTLSG 60
DB 1 MONHVISLSAAERRAHIAATFGSGRIGPQFFDALMPSERLERMAELVPGLSAHPTLSG 60
QY 61 VKACFMSAVYMEQALDGVPIYVFEEDVLLGEGAEQFLAEDTWLEERFKDSAFYVR 120
DB 61 VKACFMSAVYMEQALDGVPIYVFEEDVLLGEGAEQFLAEDTWLEERFKDSAFYVR 120
QY 121 LETMFMHVLTSVSGVADYGGRAFPILSESHCTAGYIIRKAMRFLDRFAVLPPRLHP 180
DB 121 LETMFMHVLTSVSGVADYGGRAFPILSESHCTAGYIIRKAMRFLDRFAVLPPRLHP 180
QY 121 LETMFMHVLTSVSGVADYGGRAFPILSESHCTAGYIIRKAMRFLDRFAVLPPRLHP 180
DB 121 LETMFMHVLTSVSGVADYGGRAFPILSESHCTAGYIIRKAMRFLDRFAVLPPRLHP 180
QY 181 VDLAMFGNDDEGMPVQCLNAPALCAQELHVAKFHDONSALGSLIEHRRLNRRQQRDS 240
DB 181 VDLAMFGNDDEGMPVQCLNAPALCAQELHVAKFHDONSALGSLIEHRRLNRRQQRDS 240
QY 241 PANTFKR-----LIRALTKIGREKRRQR-----EQLIGIIVPQ 279
DB 241 PANTFKR-----LIRALTKIGREKRRQR-----EQLIGIIVPQ 279
QY 234 -----RHRSRLKVMFDLKRALGFKGRKKRMRORQAELEKRYGRVILFK 280
DB 234 -----RHRSRLKVMFDLKRALGFKGRKKRMRORQAELEKRYGRVILFK 280

RESULT 13
US-08-701-191A-27
Sequence 27, Application US/08701191A
Patent No. 5942428
GENERAL INFORMATION:
APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
APPLICANT: and Steven R. Hubbard
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
OF NON-INSULIN RECEPTOR TYROSINE KINASE
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
SUITE: Suite 4700

```

```

CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/701,191A
FILING DATE: August 21, 1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Waidburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 227/088
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 304 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-701-191A-27

Query Match          5.6%; Score 82; DB 2; Length 304;
Best Local Similarity 20.8%; Pred. No. 0.39;
Matches 35; Conservative 39; Mismatches 46; Indels 48; Gaps 8;

QY 62 ERACFMSHVAVMEQALDEGVPIYVFEEDVLLGEGAEQFLAEDTWLEERFDP 113
DB 133 ERALTMGDLISFAMQISQGMQYLAEKLVHRLAANNILVAEGRKKISDFGLSRDYVE 192
QY 114 DSAFYVR-----LETMFMHVLTSVSGVADYGGRAFPILSESHCTAGYIIRKRA 162
DB 193 DS-YVRSQGRIPVKMALESLEFDHITYTQSDWSPGVLMELV-----TLGG----- 239
QY 163 MRFELDRFAVLPPRLHPYDLMVGNPDREGMPVQCLNAPALCAQELH 210
DB 240 -----NPGIIPERL--FNILKTGHRMR-----PNCSEMY 271

RESULT 14
US-09-134-001C-3799
Sequence 3799, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lyon Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
FILE REFERENCE: CTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3799
LENGTH: 337
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3799

```

Query Match 5.68; Score 81.5; DB 4; Length 337;
Best Local Similarity 22.18; Pred. No. 0.52;
Matches 50; Conservative 39; Mismatches 64; Indels 73; Gaps 13;

Search completed: May 27, 2003, 18:18:46
Job time : 30 secs

OY 1 NONHVISLASAEERAHIAATFGSCIPFOFF--DALMPSERLERAMAEIVPG---LSAH 55
Db 150 IMHILIEVDPLSKSVHLANF--LEIPFDVYPPDAIMP---LPLEADVIVGDFPIGY 204
OY 56 P-----YLSGVKACPMASHAVLMEQALD---EGVPYIAV---FEDDVLLEGAEQF 100
Db 205 PLDESRREMKLGFEKSHSYSHLLIEQISNALKGAFIVPSSHLEFDDKV--KQLENF 262
OY 101 LAEDTWLOERDPPDSAFVYVLETFMFAVLTSPSGVADYGGRAFPILSEHCCTAGYIISR 160
Db 263 IATETEMQ-----AFNLKPTLFKN-----EKARKSILLIOKKSG----- 298
OY 161 KAMRFELDRFAVLPERLHPVDLM-----FGNPDREGMPVCOLN 201
Db 299 -----ETRPVEVLIANIPDKNPOQFOGF--ISELN 327

RESULT 15

US-09-564-805-236
; Sequence 236, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 236
; LENGTH: 554
; TYPE: PRT
; ORGANISM: Synechocystis sp.
US-09-564-805-236

Query Match 5.58; Score 81; DB 4; Length 554;

Best Local Similarity 21.58; Pred. No. 1.3;
Matches 64; Conservative 37; Mismatches 99; Indels 98; Gaps 17;

OY 22 FGSRGIPF-----QFEDA-----LMPSERLERAMAEIVPGLSAHPYLSG----- 60
Db 191 YGNRLRPHRROEQFOIAETVLAAGNNILPVPPLGLAQ--EIIKILRTHHQFTGRQYN 249
OY 61 -----VEKAC-----FMSHAVL--WEQALDEGVPIYIAVFEDDVLAGE-- 95
Db 250 LMAGESVARCGDAYOGIIDLHPDNNVNFQAQHPLEFMD--DKVYPHLRPLFDD--QGELS 304
OY 96 -GAQOFLAEDTWLOERDPPDSAFVYVLETFM--HYLTSPSGVADYGGR--AFPLLESEH 150
Db 305 LSAISVIYTTTW--PALWPSPALPLGLMTVFMPLLTLPSCLVNFAWQDLEFPKYELED 362
OY 151 CGTAGYIISRKAMRFELDRFAVLPERLHPVDLMFGNPDREGMPVC-----QLN-- 201
Db 363 YLLADHSDGRNTYQLIH--LRPHL-----VYHGOPSIEDITSLSEKQSRQYLHSP 414
OY 202 -----PALCAQELHAKFHDQNSALGSLIEH--DRRLNRKQW 237
Db 415 AAGNAVALPIGDRFVQPPPPQIYEGEIHLEPNKQIHHLGEEVYIHLGQILENSRW 472

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 18:17:25 ; Search time 23 Seconds
(without alignments)
1203.008 Million cell updates/sec

Title: US-10-007-267A-8

Perfect score: 1461

Sequence: 1 MONHYISLSAASARRAHIAA.....REKRORREQLGIITVPPQ 279

Scoring table: BLOSUM62

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCF_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCFUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1461	100.0	279	12	US-10-007-267-8
2	1288	88.2	275	10	US-09-211-691-2
3	1036.5	70.9	280	12	US-10-007-267-6
4	98.5	6.7	273	9	US-10-218-381-1
5	95.5	6.5	389	9	US-10-235-056-7
6	82.5	5.6	273	9	US-09-895-913A-186
7	82.5	5.6	366	9	US-10-235-056-9
8	81	5.5	554	9	US-09-988-626-236
9	81	5.5	554	9	US-09-988-687-236
10	80	5.5	222	10	US-09-864-761-48086
11	80	5.5	285	9	US-10-253-007-54
12	80	5.5	350	9	US-10-235-056-11
13	80	5.5	363	9	US-10-235-056-11
14	80	5.5	363	10	US-09-848-021A-1
15	80	5.5	363	10	US-09-849-562A-1
16	78	5.3	302	10	US-09-815-242-5371
17	78	5.3	325	10	US-09-815-242-12608
18	77	5.3	265	9	US-09-454-279-22
19	77	5.3	570	10	US-09-764-853-551

20	76	5.2	362	9	US-10-235-056-3	Sequence 3, Appl1
21	76	5.2	755	9	US-09-738-626-4773	Sequence 4773, Ap
22	76	5.2	893	10	US-09-916-790-5	Sequence 5, Appl1
23	75	5.1	628	10	US-10-097-340-208	Sequence 208, Ap
24	74.5	5.1	275	10	US-09-841-132-443	Sequence 443, App
25	74.5	5.1	1443	10	US-09-815-242-5004	Sequence 5004, Ap
26	74.5	5.1	1451	10	US-09-815-242-10734	Sequence 10734, A
27	74	5.1	352	10	US-09-739-451-4	Sequence 4, Appl1
28	74	5.1	353	10	US-10-097-065-137	Sequence 137, App
29	74	5.1	353	9	US-09-739-451-12	Sequence 12, Appl
30	73.5	5.0	3034	10	US-09-737-149-25	Sequence 25, Appl
31	73.5	5.0	3034	10	US-09-737-149-30	Sequence 30, Appl
32	73	5.0	738	9	US-10-235-521-1	Sequence 1, Appl1
33	72.5	5.0	388	9	US-10-235-056-17	Sequence 17, Appl
34	72	4.9	3472	9	US-10-027-806-4	Sequence 4, Appl1
35	72	4.9	3472	9	US-10-034-623-4	Sequence 4, Appl1
36	72	4.9	3472	9	US-10-027-801-4	Sequence 4, Appl1
37	71.5	4.9	254	10	US-09-796-149-4	Sequence 4, Appl1
38	71.5	4.9	261	9	US-10-166-087-34	Sequence 34, Appl1
39	71.5	4.9	348	9	US-09-738-626-5823	Sequence 5823, Ap
40	71.5	4.9	436	9	US-09-764-868-826	Sequence 826, Appl
41	71.5	4.9	792	9	US-10-029-115-4	Sequence 4, Appl1
42	71.5	4.9	966	10	US-09-978-698-2	Sequence 2, Appl1
43	71.5	4.9	1244	9	US-09-789-390-13	Sequence 13, Appl
44	71.5	4.9	1273	9	US-09-789-390-11	Sequence 11, Appl
45	71.5	4.9	1295	9	US-09-789-390-30	Sequence 30, Appl

ALIGNMENTS

RESULT 1
US-10-007-267-8
Sequence 8, Application US/10007267
Patent No. US20020127682A1

GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND GENES ENCODING THEM

NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/007,267
FILING DATE: 03-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/333,412
FILING DATE: 15-Jun-1999
APPLICATION NUMBER: 08/312,387
FILING DATE: July 7, 1994

ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 279 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-007-267-8

Query Match 100.0%; Score 1461; DB 12; Length 279;
Best Local Similarity 100.0%; Pred. No. 6.2e-143;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MONHVISLAAERRAAHIAATFGSGIPFOFDALMPSERLERAAELVPGISAPHTLSG 60
DB 1 MONHVISLAAERRAAHIAATFGSGIPFOFDALMPSERLERAAELVPGISAPHTLSG 60
QY 61 VEKACFMSHAYLMQALDEGVYIAVEEDVLLGGAQFLAEDTWLQERDPPDSAFYVR 120
DB 61 VEKACFMSHAYLMQALDEGVYIAVEEDVLLGGAQFLAEDTWLQERDPPDSAFYVR 120
QY 121 LETFMHVLTPSGVADYGGRAFLPSEHSGTGAGYIISRRAMRFFLDRAFLVPPERLHP 180
DB 121 LETFMHVLTPSGVADYGGRAFLPSEHSGTGAGYIISRRAMRFFLDRAFLVPPERLHP 180
QY 181 VDLAMFGNPDREGMPVQOLPALCAOELHAKFHQNSALGSLIEHRLNRKQOMRDS 240
DB 181 VDLAMFGNPDREGMPVQOLPALCAOELHAKFHQNSALGSLIEHRLNRKQOMRDS 240
QY 241 PANTFKHRLIALTKIGRERERRRORRQOLIGKIIIVPQ 279
DB 241 PANTFKHRLIALTKIGRERERRRORRQOLIGKIIIVPQ 279

RESULT 2

US-09-211-691-2
Sequence 2, Application US/09211691
Patent No. US20020034805A1
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Young, N. Martin
APPLICANT: Wakarchuk, Warren W.
APPLICANT: National Research Council of Canada
TITLE OF INVENTION: Fusion proteins for use in Enzymatic Synthesis of
FILE REFERENCE: 019957-012910US
CURRENT APPLICATION NUMBER: US/09/211,691
CURRENT FILING DATE: 1998-12-14
PRIOR APPLICATION NUMBER: US 60/069,443
PRIOR FILING DATE: 1997-12-15
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 275
TYPE: PRT
ORGANISM: Neisseria meningitidis
US-09-211-691-2

Query Match 88.2%; Score 1288; DB 10; Length 275;
Best Local Similarity 90.0%; Pred. No. 4.6e-125;
Matches 251; Conservative 7; Mismatches 17; Indels 4; Gaps 1;

QY 1 MONHVISLAAERRAAHIAATFGSGIPFOFDALMPSERLERAAELVPGISAPHTLSG 60
DB 1 MONHVISLAAERRAAHIAATFGSGIPFOFDALMPSERLERAAELVPGISAPHTLSG 60
QY 61 VEKACFMSHAYLMQALDEGVYIAVEEDVLLGGAQFLAEDTWLQERDPPDSAFYVR 120
DB 61 VEKACFMSHAYLMQALDEGVYIAVEEDVLLGGAQFLAEDTWLQERDPPDSAFYVR 120
QY 121 LETFMHVLTPSGVADYGGRAFLPSEHSGTGAGYIISRRAMRFFLDRAFLVPPERLHP 180
DB 121 LETFMHVLTPSGVADYGGRAFLPSEHSGTGAGYIISRRAMRFFLDRAFLVPPERLHP 180
QY 181 VDLAMFGNPDREGMPVQOLPALCAOELHAKFHQNSALGSLIEHRLNRKQOMRDS 240
DB 181 VDLAMFGNPDREGMPVQOLPALCAOELHAKFHQNSALGSLIEHRLNRKQOMRDS 240

QY 241 PANTFKHRLIALTKIGRERERRRORRQOLIGKIIIVPQ 279
DB 241 PANTFKHRLIALTKIGRERERRRORRQOLIGKIIIVPQ 275

RESULT 3

US-10-007-267-6
Sequence 6, Application US/10007267
Patent No. US20020127682A1
GENERAL INFORMATION:

APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
OLIGOSACCHARIDES, AND GENES ENCODING THEM

NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:

ADDRESSEE: Klaubert & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/007,267
FILING DATE: 03-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/333,412
FILING DATE: 15-Jun-1999
APPLICATION NUMBER: 08/312,387
FILING DATE: July 7, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-007-267-6

Query Match 70.9%; Score 1036.5; DB 12; Length 280;
Best Local Similarity 71.6%; Pred. No. 4.7e-99;
Matches 209; Conservative 24; Mismatches 34; Indels 25; Gaps 3;

QY 1 MONHVISLAAERRAAHIAATFGSGIPFOFDALMPSERLERAAELVPGISAPHTLSG 60
DB 1 MONHVISLAAERRAAHIAATFGSGIPFOFDALMPSERLERAAELVPGISAPHTLSG 60
QY 61 VEKACFMSHAYLMQALDEGVYIAVEEDVLLGGAQFLAEDTWLQERDPPDSAFYVR 120
DB 61 VEKACFMSHAYLMQALDEGVYIAVEEDVLLGGAQFLAEDTWLQERDPPDSAFYVR 120
QY 121 LETFMHVLTPSGVADYGGRAFLPSEHSGTGAGYIISRRAMRFFLDRAFLVPPERLHP 180
DB 121 LETFMHVLTPSGVADYGGRAFLPSEHSGTGAGYIISRRAMRFFLDRAFLVPPERLHP 180
QY 181 VDLAMFGNPDREGMPVQOLPALCAOELHAKFHQNSALGSLIEHRLNRKQOMRDS 240
DB 181 VDLAMFGNPDREGMPVQOLPALCAOELHAKFHQNSALGSLIEHRLNRKQOMRDS 240

QY 241 PANTHER-----LIRALTKIGREKRROR-----EOLIGKIIVPQ 279
DB 234 -----RRRSRLKVMEDLKRALGKFGREKKRRERORQAELEKYGRRVILEK 280

RESULT 4

US-10-218-381-1
; Sequence 1, Application US/10218381
; Publication No. US20030013175A1
; GENERAL INFORMATION:
; APPLICANT: Satoshi KOIZUMI.
; APPLICANT: Tetsuo ENDO
; APPLICANT: Kazuhiko TABATA
; APPLICANT: Akio OZAKI
; TITLE OF INVENTION: GLYCOSYLTRANSFERASE AND DNA ENCODING THE SAME
; FILE REFERENCE: 2139.18
; CURRENT APPLICATION NUMBER: US/10/218,381
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US/09/601,519
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: PCT/JP99/00321
; PRIOR FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: JP 10-023389
; PRIOR FILING DATE: 1998-02-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 273
; TYPE: PRF
; ORGANISM: Helicobacter pylori
US-10-218-381-1

Query Match 6.7%; Score 98.5; DB 9; Length 273;
Best Local Similarity 23.9%; Pred. No. 0.042;
Matches 50; Conservative 24; Mismatches 54; Indels 81; Gaps 12;

QY 30 QFFDALMPSERLERAMAEIVPGISAHFY-----LSCVER----- 63
DB 38 QIEFDALY-KTEFGGLHPYVK-KHLHPYITONIKDMGITTNLISEVSEFYVALKYHANF 95
QY 64 -----ACFMSHAVLMEQALDEGPYIAVEFDVYLLGEGAEPLADDTWLQERFDDSAF 117
DB 96 MSIGELGCTASHYSLMEKCLELNA-FAICLEDITLKEDEKSEL--DLEKHI--QELG 149
QY 118 VVLETFMFHVLTPSPGVADYGGRAFPLESEH-----CGTAGYIISRKA 162
DB 150 YVRL-----KHLHYDPYVKE-----PLNKHNEIGERVGITKAYSHGVTGCTGYVITPKI 199
QY 163 MREFLDRAVLPPERLH-----PVDLMM 185
DB 200 AKVF-----KKHSRKWVVPVDTIM 218

RESULT 5

US-10-235-056-7
; Sequence 7, Application US/10235056
; Publication No. US20030059915A1
; GENERAL INFORMATION:
; APPLICANT: Canfield, William M.
; APPLICANT: Cummings, Richard D.
; APPLICANT: Ju, Tongzhong
; TITLE OF INVENTION: CORE 1 3-GALACTOSYL TRANSFERASES
; FILE REFERENCE: 5820.630
; CURRENT APPLICATION NUMBER: US/10/235,056
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/461,321
; PRIOR FILING DATE: 1999-12-15
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 389
; TYPE: PRF
; ORGANISM: Caenorhabditis elegans

US-10-235-056-7

Query Match 6.5%; Score 95.5; DB 9; Length 389;
Best Local Similarity 23.1%; Pred. No. 0.14;
Matches 42; Conservative 26; Mismatches 63; Indels 51; Gaps 8;

QY 2 QNHVISLASAERRAHIAATFGSGIPEOFF-----DALMPSERLERAMAEIVPGISAHFY 57
DB 115 QNH-----DKRAKHVKATWAKRCKNYVMSSEDEALPAIMLVNS-----EGRDY 159
QY 58 LSGVERACFMSHAVLMEQALDEGPYIAVEFDVYLLGEGAEPLADDTWLQERFDDSAF 117
DB 160 LMAKTGAF--KYIDHILND-----YDWFLKAD--Ddty 190
QY 118 VVLETFMFHVLTPSPGVADYGGRAFPLE-SEHCGTAGYIISRKAMRFLDRFAVLPE 176
DB 191 VVVENLRPFMLNHPDEPIHFGCKFPFPGGHSAGAVLSREALKRTIE--VALPK 248
QY 177 RL 178
DB 249 SL 250

RESULT 6

US-09-895-913A-186
; Sequence 186, Application US/09895913A
; Patent No. US20020160456A1
; GENERAL INFORMATION:
; APPLICANT: Kleantous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020160456A1el Helicobacter Polypeptides in t
; FILE REFERENCE: 06132/043002
; CURRENT APPLICATION NUMBER: US/09/895,913A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 08/881,227
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 186
; LENGTH: 273
; TYPE: PRF
; ORGANISM: Helicobacter pylori
US-09-895-913A-186

Query Match 5.6%; Score 82.5; DB 9; Length 273;
Best Local Similarity 21.4%; Pred. No. 1.9;
Matches 45; Conservative 33; Mismatches 73; Indels 59; Gaps 8;

QY 41 LERAMAEIVPGISAHFYLSGVERACFMSHAVLMEQALD--EGVPIYAVEFDVYLLGEGA 97
DB 7 LERALKE-----DLGHDLPERVLEKDFKATATVRAKQEVFGSEKY 48
QY 98 EGFLEADDTWLQ-----ERFDP-----DSAFVVLTFMFHVLTPSPGVADYGG 140
DB 49 ALLELEMTGIECYQITKDKREFKPKDALMEIRDFSLKLVETLNLNLOHSSGICATLTS 108
QY 141 RAEPLESEHCGTAGYIISRKAMRFLDRFAVLPPERLHPVDLMMGNPDRCGMPYCOL 200
DB 109 REVEALNSHKVRLDTRKTRPLRIF- EKYSVLNG-----GASHRLG-----L 151
QY 201 NPALCAQELHYAKFHQONSALGSLIERDR 230
DB 152 DDALMLKDHHLRHVKD-----LKSFLTHARK 177

RESULT 7

US-10-235-056-9
; Sequence 9, Application US/10235056


```
;; PRIOR FILING DATE: 1999-12-15
;; NUMBER OF SEQ ID NOS: 25
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 11
;; LENGTH: 350
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-235-056-11

Query Match
Best Local Similarity 26.1%; Score 80; DB 9; Length 350;
Matches 23; Conservative 14; Mismatches 19; Indels 32; Gaps 4;

QY 95 EGAEQFLAED-----TWLOERFPDPSAFVRLTETFMHVLTPSPGVADYGGRAF- 143
DB 147 EDADWFLKADDDTYIILNLRMLLSKYDPEEPI-----YFGRRFK 186

QY 144 PLESEH-CGTAGYIISRKARFFLDRF 170
DB 187 PYVKGWMSGAGYVLSKALKRFVDAT 214

RESULT 13
US-10-235-056-1
; Sequence 1, Application US/10235056
; Publication No. US2003005915A1
; GENERAL INFORMATION:
; APPLICANT: Canfield, William M.
; APPLICANT: Cummings, Richard D.
; TITLE OF INVENTION: CORE 1 3-GALACTOSYL TRANSFERASES
; FILE REFERENCE: 5820 630
; CURRENT APPLICATION NUMBER: US/10/235,056
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/461,321
; PRIOR FILING DATE: 1999-12-15
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-235-056-1

Query Match
Best Local Similarity 26.1%; Score 80; DB 9; Length 363;
Matches 23; Conservative 14; Mismatches 19; Indels 32; Gaps 4;

QY 95 EGAEQFLAED-----TWLOERFPDPSAFVRLTETFMHVLTPSPGVADYGGRAF- 143
DB 160 EDADWFLKADDDTYIILNLRMLLSKYDPEEPI-----YFGRRFK 199

QY 144 PLESEH-CGTAGYIISRKARFFLDRF 170
DB 200 PYVKGWMSGAGYVLSKALKRFVDAT 227

RESULT 14
US-09-849-031A-1
; Sequence 1, Application US/09849031A
; Patent No. US2002002603A1
; GENERAL INFORMATION:
; APPLICANT: Cummings, Richard D.
; APPLICANT: McEvel, Rodger P.
; TITLE OF INVENTION: synthetic glycosulfopeptides and methods of synthesis thereof
; FILE REFERENCE: 5820, 593
; CURRENT APPLICATION NUMBER: US/09/849,031A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 09/334,013
; PRIOR FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
```

```
;; LENGTH: 363
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-849-031A-1

Query Match
Best Local Similarity 26.1%; Score 80; DB 10; Length 363;
Matches 23; Conservative 14; Mismatches 19; Indels 32; Gaps 4;

QY 95 EGAEQFLAED-----TWLOERFPDPSAFVRLTETFMHVLTPSPGVADYGGRAF- 143
DB 160 EDADWFLKADDDTYIILNLRMLLSKYDPEEPI-----YFGRRFK 199

QY 144 PLESEH-CGTAGYIISRKARFFLDRF 170
DB 200 PYVKGWMSGAGYVLSKALKRFVDAT 227

RESULT 15
US-09-849-562A-1
; Sequence 1, Application US/09849562A
; Patent No. US20020042102A1
; GENERAL INFORMATION:
; APPLICANT: McEvel, Rodger P.
; APPLICANT: Cummings, Richard D.
; TITLE OF INVENTION: synthetic glycosulfopeptides and methods of synthesis thereof
; FILE REFERENCE: 5820, 594
; CURRENT APPLICATION NUMBER: US/09/849,562A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 09/334,013
; PRIOR FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-849-562A-1

Query Match
Best Local Similarity 26.1%; Score 80; DB 10; Length 363;
Matches 23; Conservative 14; Mismatches 19; Indels 32; Gaps 4;

QY 95 EGAEQFLAED-----TWLOERFPDPSAFVRLTETFMHVLTPSPGVADYGGRAF- 143
DB 160 EDADWFLKADDDTYIILNLRMLLSKYDPEEPI-----YFGRRFK 199

QY 144 PLESEH-CGTAGYIISRKARFFLDRF 170
DB 200 PYVKGWMSGAGYVLSKALKRFVDAT 227

Search completed: May 27, 2003, 16:23:31
Job time : 25 secs
```

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 18:12:59 ; Search time 45 Seconds
(without alignments)
596.033 Million cell updates/sec

Title: US-10-007-267a-8

Perfect score: 1461

Sequence: 1 MONHVISLASAERRAHIAA.....REKRQRREQLIGKTIIVFQ 279

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues 283224

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1329	91.0	279	2	A81971
2	1293	88.5	275	2	C81027
3	1288	88.2	275	2	S70814
4	997.5	68.3	268	2	B81971
5	928.5	63.6	280	2	A81027
6	881.5	60.3	276	2	S70815
7	461	31.6	266	2	S71025
8	418	28.6	302	1	A64077
9	271	18.5	282	1	F64091
10	231.5	15.8	268	2	AD3304
11	128	8.8	332	2	E71916
12	116.5	8.0	721	2	AH3417
13	107	7.3	738	2	T00343
14	100.5	6.9	267	2	E64175
15	99.5	6.8	273	1	B64623
16	99	6.8	492	2	T20368
17	95.5	6.5	419	2	T19837
18	93.5	6.4	444	2	F71916
19	93.5	6.4	1571	2	T14155
20	91	6.2	266	2	AG2797
21	91	6.2	266	2	H97576
22	91	6.2	735	2	S76425
23	90.5	6.2	383	2	T30730
24	90.5	6.2	404	2	C64597
25	90	6.2	284	1	E64620
26	89.5	6.1	737	2	C84232
27	89.5	6.1	1378	2	T30173
28	88	6.0	292	2	D71894
29	88	6.0	694	2	B86559

30	88	6.0	694	2	G72065	translation elonga
31	87.5	6.0	404	2	C71907	probable type II D
32	87.5	6.0	1153	2	A97179	ATP-dependent exon
33	87	6.0	256	2	T00097	hypothetical prote
34	87	6.0	273	2	E71890	probable lipopolys
35	86.5	5.9	354	2	S77119	D-alanine-D-alanin
36	86.5	5.9	845	2	T17291	hypothetical prote
37	86	5.9	258	2	AC2829	conserved hypotet
38	86	5.9	300	2	A97607	hypothetical prote
39	86	5.9	361	2	C83350	probable transcrip
40	86	5.9	5069	2	T17464	rtifamycin polyketi
41	84.5	5.8	339	2	B72072	ct389 hypothetical
42	84.5	5.8	339	2	G86551	conserved hypotet
43	84.5	5.8	410	2	G81596	protein-tyrosine k
44	83.5	5.7	402	2	B34735	protein-tyrosine k
45	83.5	5.7	435	2	JN0290	

ALIGNMENTS

RESULT 1
A81971
lacto-N-neotetraose biosynthesis glycosyl transferase NMA0525 [imported] - Neisseria
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: A81971
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491
A:Reference number: A81775; NCID:20222556; PMID:10761919
A:Accession: A81971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-279 <PAR>
A:Cross-references: GB:AL162753; GB:AL157959; NID:g7379120; PIDN:CAB83817.1; PID:g737
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: lgtB; NMA0525
C:Superfamily: lipopolysaccharide biosynthesis-associated protein

Query Match 91.0%; Score 1329; DB 2; Length 279;
Best Local Similarity 92.5%; Pred. No. 6.1e-115;
Matches 258; Conservative 5; Mismatches 16; Indels 0; Gaps 0;

QY	1	MONHVISLASAERRAHIAA	TPGSGTIPQFPDMLMSERLERMAFLVGLSAHPYLSG	60
DB	1	MONHVISLASAERRAHIAA	TPGSGTIPQFPDMLMSERLERMAFLVGLSAHPYLSG	60
QY	61	VERACFMSHAVLMEQALDEGVPIA	VEEDVLLGEGAEQFLAEDTWLQERFDPDSAEVVR	120
DB	61	VERACFMSHAVLMEQALDEGVPIA	VEEDVLLGEGAEQFLAEDTWLQERFDPDSAEVVR	120
QY	121	LETWFMHVLTPSGVADYGGRA	PPLSEHCAGTYIISRKAMFFLDRAVLPERLHP	180
DB	121	LETWFMHVLTPSGVADYGGRA	PPLSEHCAGTYIISRKAMFFLDRAVLPERLHP	180
QY	181	VDLMFENPDREMPYCOLNPAL	CAQELHYAKHDDNSALGSLTEHDDRLNRQQRDS	240
DB	181	VDLMFENPDREMPYCOLNPAL	CAQELHYAKHDDNSALGSLTEHDDRLNRQQRDS	240
QY	241	PANFMRRLRALTKIGRERER	KRRRREQLIGKTIIVFQ	279
DB	241	PANFMRRLRALTKIGRERER	KRRRREQLIGKTIIVFQ	279

RESULT 2
C81027
lacto-N-neotetraose biosynthesis glycosyl transferase lgtB NMB1928 [imported] - Neiss
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: C81027

A:Accession: A81027
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-280 <TE>
 A:Cross-references: GB:AE002541; GB:AE002098; NID:g7227175; PIDN:AAF42255.1; PID:g722718
 A:Experimental source: serogroup B, strain MC58
 C:Genetics:
 A:Gene: NMB1926
 C:Superfamily: lipopolysaccharide biosynthesis-associated protein

Query Match 63.6%; Score 928.5; DB 2; Length 280;
 Best Local Similarity 65.4%; Pred. No. 5.6e-78;
 Matches 191; Conservative 28; Mismatches 46; Indels 25; Gaps 3;

QY 1 MONHVISLAAERRAHIAATFGSGRIPGFEDALMPSERLERAAEIVPGISAHPIYLSG 60
 |||||
 DB 1 MONHVISLAAERRAHIAATFGVGRIPOFEDALMPSELNRMAAEIVPGIAKHLLSE 60
 QY 61 VEKACFMSHAIVLMKQALDEGVYIAVFEDDVLGSGAEQFLAEDTWLOERPPDSAFVVR 120
 |||||
 DB 61 VEKACFMSHAIVLMKQALDEGLPYVAVFEDDVLGKDAEFLAEDTWLEERFPKDSAFIVR 120
 QY 121 LETMFAMVLTSPGVADYGGRAFPLESEHCGTAGIISRKAMREFLDRAVLPPERLHP 180
 |||||
 DB 121 LETMFAMVLTSPGVADYGGRAFPLESEHCGTAGIISRKAMREFLDRAVLPPERLHP 180
 QY 181 VDLMEFGNPDDEGMPVQOLNPAALCAQELHYAKFHDONSALGSLIEHNRRLNRKQWRDS 240
 |||||
 DB 181 VDLMEFGNPDDEGMPVQVNPALCTQELHYAKFLSKNSMGSDEKDEKQGR----- 233
 QY 241 PANTFKR-----LIRALTKIGRERKROR-----EOLIGKIIVPQ 279
 |||||
 DB 234 -----RHRSLSKVMEDLKRALGKFGREKKRMEROAROALEKAYGRVYSFK 280

RESULT 6

S70815
 glycosyl transferase E (EC 2.4.-.-) - Neisseria meningitidis
 C:Species: Neisseria meningitidis
 C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 20-Sep-1999
 C:Accession: S70815
 R:Jennings, M.P.; Hood, D.W.; Peak, I.R.A.; Vitrj, M.; Moxon, E.R.
 MOL. Microbiol. 16, 729-740, 1995
 A:Title: Molecular analysis of a locus for the biosynthesis and phase-variable expression
 A:Reference number: S70812; MUID:96414473; PMID:8817494
 A:Accession: S70815
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-276 <JEN>
 A:Cross-references: EMBL:U25839; NID:g973183; PIDN:AAC44086.1; PID:g973187
 A:Gene: lytE
 C:Superfamily: lipopolysaccharide biosynthesis-associated protein
 C:Keywords: glycosyltransferase

Query Match 60.3%; Score 881.5; DB 2; Length 276;
 Best Local Similarity 63.4%; Pred. No. 1.2e-73;
 Matches 185; Conservative 27; Mismatches 51; Indels 29; Gaps 4;

QY 1 MONHVISLAAERRAHIAATFGSGRIPGFEDALMPSERLERAAEIVPGISAHPIYLSG 60
 |||||
 DB 1 MONHVISLAAERRAHIAATFGVGRIPOFEDALMPSELNRMAAEIVPGIAKHLLSE 60
 QY 61 VEKACFMSHAIVLMKQALDEGVYIAVFEDDVLGSGAEQFLAEDTWLOERPPDSAFVVR 120
 |||||
 DB 61 VEKACFMSHAIVLMKQALDEGLPYVAVFEDDVLGKDAEFLAEDTWLEERFPKDSAFIVR 116
 QY 121 LETMFAMVLTSPGVADYGGRAFPLESEHCGTAGIISRKAMREFLDRAVLPPERLHP 180
 |||||
 DB 121 LETMFAMVLTSPGVADYGGRAFPLESEHCGTAGIISRKAMREFLDRAVLPPERLHP 180
 QY 181 VDLMEFGNPDDEGMPVQOLNPAALCAQELHYAKFHDONSALGSLIEHNRRLNRKQWRDS 240
 |||||

DB 177 VDLMEFGNPDDEGMPVQVNPALCTQELHYAKFLSKNSMGSDEKDEKQGR----- 229
 QY 241 PANTFKR-----LIRALTKIGRERKROR-----EOLIGKIIVPQ 279
 |||||
 DB 230 -----RHRSLSKVMEDLKRALGKFGREKKRMEROAROALEKAYGRVYSFK 276

RESULT 7

S71025

lipopolysaccharide biosynthesis protein lfc2B - Haemophilus influenzae (strain RM7004

C:Species: Haemophilus influenzae
 C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
 C:Accession: S71025

R:High, N.J.; Jennings, M.P.; Moxon, E.R.
 MOL. Microbiol. 20, 165-174, 1996

A:Title: Random repeats of the tetramer 5'-CAAT-3' present in lfc2A are required for
 A:Reference number: S71024; MUID:97014379; PMID:8861214

A:Accession: S71025

A:Molecule type: DNA

A:Residues: 1-266 <HIC>

A:Cross-references: EMBL:Z54182; NID:g1150402; PIDN:CA90892.1; PID:g1150403
 A:Experimental source: strain RM7004

A:Gene: lfc2B

A:Function:

A:Description: required for the biosynthesis of a phase-variable lipopolysaccharide s
 C:Superfamily: lipopolysaccharide biosynthesis-associated protein

Query Match 31.6%; Score 461; DB 2; Length 266;
 Best Local Similarity 39.1%; Pred. No. 7.2e-35;
 Matches 109; Conservative 46; Mismatches 106; Indels 18; Gaps 6;

QY 3 NHVISLAAERRAHIAATFGSGRIPGFEDALMPSERLERAAEIVPGISAHPIYLSGVE 62
 |||||
 DB 4 NVVISLTAQYQREHAKQKESQONIDFEEDALKPSEKELSLIEKRIPLV-LHAKITGE 62
 QY 63 KACFMSHAIVLMKQALDEGVYIAVFEDDVLGSGAEQFLAEDTWLOERPPDSAFVVRLE 122
 |||||
 DB 63 KACFMSHAIVLMKQALDEGVYIAVFEDDVLGSGAEQFLAEDTWLOERPPDSAFVVRLE 122
 QY 123 TMEFAMVLTSPGVADYGGRAFPLESEHCGTAGIISRKAMREFLDRAVLPPERLHPVD 182
 |||||
 DB 123 TMEFAMVLTSPGVADYGGRAFPLESEHCGTAGIISRKAMREFLDRAVLPPERLHPVD 182
 QY 183 LMEFGNPDDEGMPVQOLNPAALCAQELHYAKFHDONSALGSLIEHNRRLNRKQWRDS 240
 |||||
 DB 183 LMEFGNPDDEGMPVQVNPALCTQELHYAKFLSKNSMGSDEKDEKQGR----- 231
 QY 241 PANTFKR-----LIRALTKIGRERKROR-----EOLIGKIIVPQ 279
 |||||
 DB 232 -----RHRSLSKVMEDLKRALGKFGREKKRMEROAROALEKAYGRVYSFK 266

RESULT 8

A64077

lipopolysaccharide biosynthesis protein lfc2A - Haemophilus influenzae

C:Species: Haemophilus influenzae
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 28-Jul-2000
 C:Accession: A64077; S15282; S39576

R:Feilschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kellavag
 A:Author: Gocayne, J.D.; Scott, J.; Shiley, R.; Liu, L.T.; Glodok, A.; Kelley, J.M.; Weidman
 A:Author: Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghegan, N.S.M.
 Science 269, 496-512, 1995

A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:9550630; PMID:7542800

A:Accession: A64077

A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA

A:Residues: 1-302 <TIG>

A:Cross-references: GB:U32736; GB:I42023; NID:g1573519; PIDN:AAC22208.1; PID:g1573535
 A:Experimental source: strain Rd KM20
 R:Cooper, L.D.; Yocum, R.; Mertsola, J.; Latimer, J.L.; Hanson, M.S.; McCracken Jr., G.

Query Match	8.8%;	Score 128;	DB 2;	Length 332;
Best Local Similarity	23.8%;	Pred. No. 0.00051;		
Matches 77; Conservative	36;	Mismatches 108;	Indels 102;	Gaps 17;

```

RESULT 12
AH3417
Ipsa protein [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AH3417
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzei, P.H.; Hagius, S.; O'Callaghan, D.; Letessier
.; Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A>Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AH3417
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1721 <KMB>
A:Cross-References: GB:AE008917; PIDN:AL52507.1; PID:gi7983318; GSPDB:GN00190
A:Experimental source: Strain 16M
C:Genetics:
A:Gene: BMel1326
A:Map position: 1

```

Query Match	8.0%;	Score 116.5;	DB 2;	Length 721;
Best Local Similarity	25.8%;	Pred. No. 0.016;		

RESULT 1
T00343

Query Match	7.3%	Score 107,	DB 2,	length 738;
Best Local Similarity	23.3%	Pred No. 0.13,		
Matches 44, Conservative	40,	Mismatches 83,	Indels 22,	Gaps 7,

Db 62

RESULT 14
 E64175
 hypothetical protein H11697 (1sg locus) - Haemophilus influenzae
 C1Species: Haemophilus influenzae
 C1Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 08-Oct-1999
 C1Accession: E64175; S27580
 R.Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
 ; Gocayne, J.D.; Scott, J.; Sherry, R.; Liu, L.I.; Glodde, A.; Kelley, J.M.; Weidman
 ; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fulbright, J.L.; Geoghegan, N.S.M.
 Science 269, 496-512, 1995
 A1Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter
 A1Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A1Reference number: A64000; MUID:95350630; PMID:7542800
 A1Accession: E64175

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 18:06:19 ; Search time 11 seconds
(without alignments)
1051.991 Million cell updates/sec

Title: US-10-007-267a-8

Perfect score: 1461
Sequence: 1 MONHVISLASAERRAHIAA.....REKRORREQLIGKIVPEQ 279

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1461	100.0	279	1 LGTB_NEIGO	050947 neisseria g
2	1339	91.0	279	1 LGTB_NEIMA	P57033 neisseria m
3	1293	88.5	275	1 LGTB_NEIMB	051116 neisseria m
4	1036.5	70.9	280	1 LGTB_NEIGO	050950 neisseria g
5	928.5	63.6	280	1 LGTB_NEIMB	051117 neisseria m
6	461	31.6	266	1 LGTB_HAEMIN	057394 haemophilus
7	418	28.6	302	1 LGTB_HAEMIN	003974 haemophilus
8	321	22.0	263	1 LGTB_PASHA	005770 pasteurella
9	271	18.5	282	1 LGTB_HAEMIN	057125 haemophilus
10	100.5	6.9	257	1 LGTB_HAEMIN	P71398 haemophilus
11	88	6.0	694	1 EFG_CHLPR	092802 chlamydia p
12	86.5	5.9	354	1 DDG_SYNY3	P76332 synecocyst
13	83.5	5.7	741	1 PIFA_ECOLI	P96329 escherichia
14	83.5	5.7	971	1 KBLI_MOUSE	P25799 mus musculu
15	83.5	5.7	1114	1 RETI_HUMAN	P07949 homo sapien
16	82.5	5.6	273	1 NADC_HELPY	025909 helicobacte
17	82	5.6	1045	1 RAGI_XENLA	091829 xenopus lae
18	81	5.5	275	1 HUPH_RHOCA	003007 rhodobacter
19	81	5.5	380	1 CAIA_SALTY	082422 salmonella
20	81	5.5	554	1 Y514_SYNY3	055470 synecocyst
21	81	5.5	726	1 RNR_MYCPN	P75529 mycoplasma
22	81	5.5	790	1 ATST_SYNY7	P37385 synecococc
23	80.5	5.5	429	1 MYCS_RAT	P23999 rattus norv
24	80	5.5	1237	1 B3A2_RABIT	P48746 oryctolagus
25	79.5	5.4	462	1 TRPE_LEPBI	P20463 leprospira
26	79.5	5.4	968	1 KBLI_HUMAN	P19988 homo sapien
27	79	5.4	522	1 KBLI_RAT	063339 rattus norv
28	79	5.4	933	1 GLND_RATLO	098c27 thizobium l
29	78.5	5.4	344	1 KAND_SCHPO	010078 schizosacch
30	78.5	5.4	1268	1 YRDE_CAEEL	009575 caenorhabdi
31	78	5.3	395	1 Y130_ARCFU	030107 archaeoglob
32	78	5.3	395	1 LDOX_MAIZE	P41223 zea mays (m
33	78	5.3	424	1 CRT3_ARATH	004153 arabidopsis

34	78	5.3	1234	1 B3A2_RAT	P2347 rattus norv
35	78	5.3	1237	1 B3A2_MOUSE	P13808 mus musculu
36	78	5.3	1238	1 B3A2_CAVPO	092098 cavia porce
37	78	5.3	1241	1 B3A2_HUMAN	P04920 homo sapien
38	77.5	5.3	273	1 NADC_HELPY	092312 helicobacte
39	77.5	5.3	380	1 CAIA_SALTY	082912 salmonella
40	77.5	5.3	965	1 CAPP_SOLTU	P29196 solanum tub
41	77	5.3	694	1 EFG_CHLTR	084444 chlamydia t
42	76.5	5.2	694	1 EFG_CHLMU	09136 chlamydia m
43	76.5	5.2	967	1 CAPP_PEA	P51062 pisum sativ
44	76.5	5.2	1115	1 RET_MOUSE	P35546 mus musculu
45	76	5.2	949	1 GLND_RHIME	P56884 thizobium m

ALIGNMENTS

RESULT 1
ID LGTB_NEIGO STANDARD; PRT; 279 AA.

AC 050947;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lacto-N-neotetraose biosynthesis glycosyl transferase lgtb.
GN LGTB.
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID:485;

RA [1]
RC STRAIN=F62;
RX MEDLINE=95053752; PubMed=7964493;
RA Gotschlich E.C.;
RT "genetic locus for the biosynthesis of the variable portion of
RT Neisseria gonorrhoeae lipooligosaccharide.";
RL J. Exp. Med. 180:2181-2190(1994).
CC -!- FUNCTION: ADDS THE SECOND GALACTOSE TO THE LACTO-N-TETRAOSE CHAIN
CC IN LOS.
CC -!- PATHWAY: BIOSYNTHESIS OF THE TERMINAL LACTO-N-NEOTETRAOSE LPS
CC STRUCTURE.

CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 25.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: U14554; AAA68010.1; -
DR InterPro: IPR002654; GT_25.
DR Pfam: PF01755; Glyco_transf_25; 1.
KW Lipopolysaccharide biosynthesis; Transferase; Glycosyltransferase.
SQ SSQUNCE 279 AA; C2219DIA6119D622 CRC64;

Query Match 100.0%; Score 1461; DB 1; Length 279;
Best Local Similarity 100.0%; Pred. No. 6.9e-128;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MONHVISLASAERRAHIAAIFGSGRIPFQFDALMPSERLERAMAEIVPGLSNHPLSG	60
DB	1	MONHVISLASAERRAHIAAIFGSGRIPFQFDALMPSERLERAMAEIVPGLSNHPLSG	60
QY	61	VERKCEFSHAVLWQALDEGVPIAIVFEDVYLGEAGBOFLAETVQERDPDSATVVR	120
DB	61	VERKCEFSHAVLWQALDEGVPIAIVFEDVYLGEAGBOFLAETVQERDPDSATVVR	120
QY	121	LETFEMVILVSPSGVADYCGRAPFLSEHCCTAGYIISRKAMFFLDRAVLPPEPLHP	180
DB	121	LETFEMVILVSPSGVADYCGRAPFLSEHCCTAGYIISRKAMFFLDRAVLPPEPLHP	180

```
OY 181 VDLMEFGNPDDEGMPVQOLNPAALCAQELHYAKFHNDONSALGSLIEHNRRLNRKQWRDS 240
DB 181 VDLMEFGNPDDEGMPVQOLNPAALCAQELHYAKFHNDONSALGSLIEHNRRLNRKQWRDS 240
OY 241 PANTFKHRLIRALTKIGRERERKRRORRROLGKIITVPPQ 279
DB 241 PANTFKHRLIRALTKIGRERERKRRORRROLGKIITVPPQ 279

RESULT 2
ID LGTB_NEIMA STANDARD: PRT: 279 AA.
AC P57033;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lacto-N-neotetraose blosynthesis glycosyl transferase lgtrB.
GN LGTB OR NMA0525.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Z2491 / Serogroup A / Serotype 4A;
RX MEDLINE-20222556; PubMed-10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Felwell T., Hamlin N., Holroyd S.,
RA Jagers K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis Z2491."
RL Nature 404:502-506(2000).
-1- FUNCTION: ADDS THE SECOND GALACTOSE TO THE LACTO-N-TETRAOSE CHAIN
IN LOS.
-1- PATHWAY: BIOSYNTHESIS OF THE TERMINAL LACTO-N-NEOTETRAOSE LPS
STRUCTURE.
-1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 25.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: AL162753; CAB83817.1; -
DR InterPro: IPR002654; GT_25;
DR Pfam: PF01753; Glyco_transf_25; 1.
KW Lipopolysaccharide biosynthesis; Transferase; Glycosyltransferase;
KW Complete proteome.
SQ SEQUENCE 279 AA; 31903 MW; 8703B56513A0D347 CMC64;

Query Match 91.0%; Score 1329; DB 1; Length 279;
Best Local Similarity 92.5%; Pred. No. 1.le-115;
Matches 258; Conservative 5; Mismatches 16; Indels 0; Gaps 0;
```

```
DB 181 VDLMEFGNPDDEGMPVQOLNPAALCAQELHYAKFHNDONSALGSLIEHNRRLNRKQWRDS 240
OY 241 PANTFKHRLIRALTKIGRERERKRRORRROLGKIITVPPQ 279
DB 241 PANTFKHRLIRALTKIGRERERKRRORRROLGKIITVPPQ 279

RESULT 3
ID LGTB_NEIMA STANDARD: PRT: 275 AA.
AC 051116;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lacto-N-neotetraose blosynthesis glycosyl transferase lgtrB.
GN LGTB OR NMB1928.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MC58 / Serogroup B;
RX MEDLINE-96414473; PubMed-8817494;
RA Jennings M.P., Hood D., Peak I.R.A., Virji M., Moxon E.R.;
RT "Molecular analysis of a locus for the biosynthesis and phase-variable
RT expression of the Lacto-N-neotetraose terminal lipopolysaccharide
RT structure in Neisseria meningitidis."
RL Mol. Microbiol. 18:729-740(1995).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-MC58 / Serogroup B;
RX MEDLINE-20175755; PubMed-10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Clecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Yamathavan J.,
RA Gill J., Scarlato V., Maignani V., Pizsa M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappapoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58."
RL Science 287:1809-1815(2000).
-1- FUNCTION: ADDS THE SECOND GALACTOSE TO THE LACTO-N-TETRAOSE CHAIN
IN LOS.
-1- PATHWAY: BIOSYNTHESIS OF THE TERMINAL LACTO-N-NEOTETRAOSE LPS
STRUCTURE.
-1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 25.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: U25839; AAC44085.1; -
DR EMBL: AE002541; AAF42257.1; -
DR TIGR: NMB1928; -
DR InterPro: IPR002654; GT_25;
DR Pfam: PF01753; Glyco_transf_25; 1.
KW Lipopolysaccharide biosynthesis; Transferase; Glycosyltransferase;
KW Complete proteome.
FT CONFLICT 97 A -> E (IN REF. 1).
SQ SEQUENCE 275 AA; 31578 MW; E871305E2F6CF70F CMC64;

Query Match 88.5%; Score 1293; DB 1; Length 275;
Best Local Similarity 90.3%; Pred. No. 2.4e-112;
Matches 252; Conservative 7; Mismatches 16; Indels 4; Gaps 1;
```

Db	1	MQNHVISLASAAERRAHIAIDTFGRHGIPIQDFDIALMPSERLEQOMALFVGLSAPHYISG	60
Qy	61	VERACFSMHNATMCAIDDECYPTIYAVVEDVYLGECAEOPFLAEDTWLOEFPDDSAFYVR	120
Db	61	VERACFSMHNATMCAIDDECYPTIYAVVEDVYLGECAEOPFLAEDTWLOEFPDDPTAFYVR	120
Qy	121	LETMFMHVLTSPSGVADYGGAFFLLESEHGCTAGYIISKRAMFFLDRAVFLPERLHP	180
Db	121	LETMFMHVLTSPSGVADYCGGAFFLLESEHGCTAGYIISKRAMFFLDRAVFLPERLHP	180
Qy	181	VDLMMFGNPDDEBGMPCOLNLPALCAOELHTAKFHONSALGSLIEHDLRLNKKQWRDS	240
Db	181	VDLMMFSDFEPRBGMPCOLNLPALCAOELHTAKFHONSALGSLIEHDLRLNKKQRDS	240
Qy	241	PANTFKRLRLALTKRIGEREKREKQREBOILGKTIIVPQ	279
Db	241	PANTFKRLRLALTKRISEREKRRQREBO-----FIVPQ	275

```

RESULT 4
LGTE_NEIGO LGTE_NEIGO STANDARD: PRT: 280 AA.
AC Q50930;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lacto-N-neotetraose biosynthesis glycosyl transferase lgte.
GN LGTE.
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=F62;
RA MEDLINE=95053752; PubMed=7964493;
RA Gotschlich E.C.;
RT "Genetic locus for the biosynthesis of the variable portion of
RL Neisseria gonorrhoeae lipooligosaccharide." Dec1994.
J. Exp. Med. 180:2181-2190(1994).
-1 FUNCTION: ADOS THE FIRST GALACTOSE TO THE LACTO-N-TETRAOSE CHAIN
IN LOS.
-1 PATHWAY: BIOSYNTHESIS OF THE TERMINAL LACTO-N-NEOTETRAOSE LPS
STRUCTURE.
-1 SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 25.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/isb-sib.ch).
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U14554; AAA68013.1; -.
DR InterPro: IPR002654; GT_25.
DR Pfam: PF01753; Glyco_transf_25; 1.
KW Lipopolysaccharide biosynthesis; Transferase; Glycosyltransferase.
SQ SEQUENCE 280 AA; 32420 MW; F043597BA3040407 CRC64;

Query Match 70.9%; Score 1036.5; DB 1; Length 280;
Best Local Similarity 71.6%; Pred. No. 1.3e-88;
Matches 209; Conservative 24; Mismatches 34; Indels 25; Gaps 3.

OY 1 MONNHTSLASAERRAHIAATGSGRIPQFPDAMPSERLRRAMAELVPGSAHPYLSG 60
Db 1 MONNHTSLASAERRAHIAADTGTSGKIRIPQFDALMPSERLQAAAEVPGSAHPYLSG 60
OY 61 VERACFMSHAVIMEQALDEGVPIYVFEEDVVLIGGAEQFLAEDWTLOERFPDPSAFVVR 120
Db 61 VERACFMSHAVIMEQALDEGLPIYVAFEEEDVVLIGGAEQFLAEDWTLPLEERFPKDSAFIVR 120
OY 121 LETMRAHVITSGVADYGGRAFPLLSESHCGTAGTIIIRKMRFFLDFNVAVLPERRLHP 180

```

Db	121	LEIEMAKTIVAPDKLVANENSSFLBSEHOGTAGYIISREAMREFLDRAVLPERIKA	180
Oy	161	VOLMEFGPDDRECHPCVQOLPALCAOELHTAKFDONSALGSLIENDRLNRKQOMRDS	240
Db	161	VOLMEFTYFPEKGMPIYQVSPALCTOELHYAKFELSONSMGSDLEKDEOGR	233
Oy	241	PANFKHR-----LIRALTGKIREKREKROR-----BOLIKILVPRO	279
Db	234	-----RHRSLKVMFDLKRALCKFEREKKKRNRORQALEKVGIRVILTK	280
RESULT 5			
LGTE_NEIMB	STANDARD:	PRT:	280 AA.
AC	051117;		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Lacto-N-neotetraose blosynthesis glycosyl transferase lgte.		
GN	LGTE OR NMB1926.		
OS	Neisseria meningitidis (serogroup B).		
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.		
OX	NCBI_TaxID=491;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-MC58 / Serogroup B:		
RX	MEDLINE=96414473; PubMed=8817494;		
RA	Jennings M.P., Hood D., Peak I.R.A., Virji M., Moxon E.R.;		
RT	"Molecular analysis of a locus for the biosynthesis and phase-variable		
RT	expression of the lacto-N-neotetraose terminal lipopolysaccharide		
RL	structure in Neisseria meningitidis.";		
RL	Mol. Microbiol. 18:729-740(1995).		
RP	[2]		
RN	SEQUENCE FROM N.A.		
RC	STRAIN-MC58 / Serogroup B:		
RX	MEDLINE=20175755; PubMed=10710307;		
RA	Tettelein H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,		
RA	Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,		
RA	Nelson W.C., Gwyn M.L., Deroy R., Peterson J.D., Hickey E.K.,		
RA	Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,		
RA	Mason T.C., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,		
RA	Cotton M.D., Uterback T.R., Khouri H., Qin H., Yamachyan J.,		
RA	Gill J., Scarlato V., Masignani V., Pizze M., Grandi G., Sun L.,		
RA	Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;		
RT	"Complete genome sequence of Neisseria meningitidis serogroup B strain		
RT	MC58.";		
RL	Science, 287:1809-1815(2000).		
CC	-1- FUNCTION: ADSR THE FIRST GALACTOSE TO THE LACTO-N-TETRAOSE CHAIN		
CC	IN LOS.		
CC	-1- PATHWAY: BIOSYNTHESIS OF THE TERMINAL LACTO-N-NEOTETRAOSE LPS		
CC	STRUCTURE.		
CC	-1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 25.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; U25839; AAC44086.1; -		
DR	EMBL; AE002541; AAF42255.1; -		
DR	TIGR; NMB1926; -		
DR	InterPro; IPR002654; GT_25.		
DR	Pfam; PF01755; Glyco_transf_25; 1.		
KW	Lipopolysaccharide biosynthesis; Transferase; Glycosyltransferase;		
KW	complete proteome.		
FT	CONFLICT 16 A -> G (IN REF. 1).		
FT	CONFLICT 89 92 MISSING (IN REF. 1).		
FT	CONFLICT 176 178 EMI -> RVD (IN REF. 1).		
SO	SEQUENCE 280 AA; 32790 MW; 826827E942B5842 CRC64;		


```

QY 1 MONHYISLASAAERRAHIAATFGSGNGIDPFQFDALMPSERLERAAAEIVPGLSAHPYISG 60
Db 1 MONHYISLASAAERRAHIAATFGVNGIDPFQFDALMPSEELNRMAAEIVPGLAKOHLISE 60
QY 61 VEKAEFMSHAIVMEQALDEGVPIAVFEDDYVLGGSAQOFLAEDYVWLOERPDPSAFYVR 120
Db 61 VEKAEFMSHAIVMEQALDEGLPIAVFEDDYVLGGDAEFLAEDYVLEERFVKDSAFYVR 120
QY 121 LETMEHVIYLTSPSGVADYGGRAFPPLISESHGCTAGYIISRKRAAEFFLDRAVLPERLHP 180
Db 121 LETMEHVIYVRDXYLVNENRSEFPLISESHGCTAGYIISREARAEFFLEFVAVLPEWIK 180
QY 181 VDLMEFGNPDDEGKMPVQCLNPAACAQELHYAKFDONSALGSLIEHNRRLNRKQWRDS 240
Db 181 VDMMEFTYFEDEKMPVQVNPALCTOELHYAKFPLSKNSMGLSDSEKREOER----- 233
QY 241 PANTFKHR-----LIRALFKIRERERKRQR-----EQLGKIITVPO 279
Db 234 -----RHRSLKVMEDLKRALGKFERREKKRERORQALEERAYGRVYSFK 280

RESULT 6
L12B_HAEIN STANDARD: PRT; 266 AA.
AC 057394;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE LipoOligosaccharide biosynthesis protein l1c2b.
GN l1c2b.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
CC Haemophilus.
CC NCB1_TaxID=727;
RN: [1]
RP SEQUENCE FROM N.A.
RA STRAIN=RM 7004 / Serotype B;
RA High N.J., Jennings M.P., Moxon R.E.;
RL Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: INVOLVED IN EXTRACELLULAR LIPOOLIGOSACCHARIDES (LOS)
CC BIOSYNTHESIS AND VIRULENCE EXPRESSION. INVOLVED IN THE SYNTHESIS
CC OF THE OLIGOSACCHARIDE MOIETY OF THE LOS MOLECULE BY ADDING GALAC
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 25.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z54182; CA90892.1; -
DR EMBL; U36398; AAA84947.1; -
DR InterPro; IPR002654; GT_25.
DR Pfam; PF01755; Glyco_transf_25; 1.
RV Transferase; Glycosyltransferase; Virulence.
SO SEQUENCE 266 AA; 31131 MW; 0C68AE9BA249BD39 CRC64;

Query Match 31.6%; Score 461; DB 1; Length 266;
Best Local Similarity 39.1%; Pred. No. 2.1e-35;
Matches 109; Conservative 46; Mismatches 106; Indels 18; Gaps 6;

QY 3 NHVITSLASAEERRAIIATFGSGNIPFQFDALMPSERLERAAAEIVPLSAHPYISGE 62
Db 4 NVVITSLATYQREHHIOKEFSGOONIPFEFDLAKRSKELTSLIEKFIPLV-LHAKITBEE 62

63 KACFMSHAIVMEQALDEGVPIAVFEDDYVLGGSAQOFLAEDYVWLOERPDPSAFVRL 122

```

D6 KACPFISHMTLMKCCSEDDLPYIIVTIEDVVLGENDAKLAEDEWLEAFKQCDRTILAFE 122
D7 123 TWFMHVLTSPSGADVADYGGAFLPLESEHCCTAGATIIISKRAMEFLDRPVALPPERLAHVD 182
D8 123 TPLNFSKCCKDKIKRPSYSGRIKLKTLSVENCNGAAYISEAVNKOLSAHTCSITSMLIAID 182
D9 103 LMMFENPDREGMPYCOPALPCAOELHYAKRHODNSALGSIEHDR- -LNMRKOOWDS 240
D10 103 LMF-----NIHQSYOVSPEGVCE---GGLYPKDIKHSLTEFEROKITSVKK---- 231

Oy 241 PANTFKRLIRIALTKIGREKERORROBQLGIKIIVPO 279
Db 232 --RTLKVTLISLAG--PRKILRKITYRLKLFISKHIIVPR 266

RESULT 7
LEX1_HAEIN STANDARD; PRF; 302 AA.
AC 003974; Q48209;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE lipooligosaccharide biosynthesis protein lex-1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
NCBI_TaxId=727;

[1]
RN SEQUENCE FROM N.A.
RP STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McEweney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shiley R., Liu L.-I., Giodel A., Kelley J.K.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Usterback T.R., Hanna M.C., Nguyen D.T., Saudke D.M., Brandon R.C.,
RA Fine L.D., Fritcham J.L., Furumann J.L., Georgagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RD Rd."; Science 269:496-512(1995).

[2]
RN SEQUENCE FROM N.A.
RP STRAIN-DIA2 / Serotype B;
RX MEDLINE=92065807; PubMed=1956289;
RA Cope I.D., Yegorov R., Mertsolea J., Latimer J.L., Hanson M.S.,
RA McCracken G.H. Jr., Hansen E.J.;
RT "Molecular cloning of a gene involved in lipooligosaccharide
biosynthesis and virulence expression by Haemophilus influenzae type
B"; Mol. Microbiol. 5:1113-1124(1991).

[3]
RN SEQUENCE FROM N.A.
RP STRAIN-RM 7004 / Serotype B;
RX MEDLINE=95020659; Pubmed=752834;
RA High N.J., Deedman M.E., Monox E.R.;
RT "The role of a repetitive DNA motif (5'-CAAT-3') in the variable
expression of the Haemophilus influenzae lipopolysaccharide epitope
alpha gal(1-4)beta gal."; Mol. Microbiol. 9:1275-1282(1993)

-1- FUNCTION: INVOLVED IN EXTRACELLULAR LIPOOLIGOSACCHARIDES (LOS)
BIOSYNTHESIS AND VIRULENCE EXPRESSION. INVOLVED IN THE SYNTHESIS
OF THE OLIGOSACCHARIDE MOETTY OF THE LOS MOLECULE BY ADDING GALAC.
-1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 25.
-1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-11 IS THE INITIATOR.

-- THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL: U32736; AAC22208.1; -
DR EMBL: X56903; CAA0221.1; -
DR EMBL: L19441; AAB65534.1; -
DR PIR: S15282; S15282.
DR TIGR: H10550; -
DR InterPro: IPR002654; GL_25.
DR Pfam: PF01755; Glyco_transf_25; 1.
KW Transferase; Glycosyltransferase; Virulence; Complete proteome.
FT VARIANT 29 K -> N (IN STRAINS DL42 AND RM 7004).
FT VARIANT 63 MISSING (IN STRAIN RM 7004).
FT VARIANT 70 MISSING (IN STRAIN DL42).
FT VARIANT 106 F -> L (IN STRAINS DL42 AND RM 7004).
FT VARIANT 151 Q -> R (IN STRAINS DL42 AND RM 7004).
FT VARIANT 256 S -> P (IN STRAIN RM 7004).
FT VARIANT 274 N -> D (IN STRAIN RM 7004).
SQ SEQUENCE 302 AA; 35490 MW; BFC204F9B3372D2C CRC64;

Query Match 28.6%; Score 418; DB 1; Length 302;
Best Local Similarity 32.9%; Pred. No. 2,4e-31;
Matches 100; Conservative 63; Mismatches 109; Indels 32; Gaps 6;

OY 1 MONHVISLASAERRAHIAATFGSRGIPFOFDAL-----MPSERL 41
DB 4 IENVIVSMENATERRKHITKOFESKLSFFFNATYYTOSINOSINOSINOSINOSI 63
OY 42 ERMAE---LVPLSAHPVYISGVKACFMSHVAWQALDEGVPIYAVFEDVLLSGAE 98
DB 64 NOSINOSINSHIESRILITKGECLISHFLMKNKCVNENFEYKIFEDDVLIGENAE 123
OY 99 OFLAEDTWLOERFDPDSAFVVRLETMEPMY-LTSPSGVADYGRAPFLISEHCGTAGY 157
DB 124 VFLNQMENKTRDPNDIPRIETFLQPKLEKQIKIPFNRRNDILKSTWGTAGYI 183
OY 158 ISKRAEFLDRFVLPPELHPVDLMFEGNPDREGMPYQALPALCAQELHYAFHQ 217
DB 184 ISGAKAYVEYELKIPSEIYVADELIFNKLVADVNYIYQALPAICQEL---QANQS 240
OY 218 NSALGSLIEHRLN---RQOWRDSFANTFKRLIRALTKIGREKRRQREOLIGKI 274
DB 241 KSVLTSGLEKERRKSKIRKKTLLKORLFRIRKENIIRAL---NRKWKKEORIKEMQKE 297
OY 275 IYVF 278
DB 298 IYVF 301

RESULT 8
LPSA_PASHA
ID LPSA_PASHA STANDARD; PRT; 263 AA.

AC 005770; 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE LPSA protein.
GN LPSA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Mannheimia.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Serotype A1;
RX MEDLINE=95301198; PubMed=7781993;
RA Potter M.D., Lo R.Y.C.;
RT Cloning and characterization of a gene from Pasteurella haemolytica
RT A1 involved in lipopolysaccharide biosynthesis.*;
RL FEMS Microbiol. Lett. 129:75-81(1995).
RN [2]

RP SEQUENCE OF 189-263 FROM N.A.

RC STRAIN-Serotype A1;
RX MEDLINE=91358346; PubMed=1885539;
RA Abdullah K.M., Lo R.Y.C., Mellors A.;
RT Cloning, nucleotide sequence, and expression of the Pasteurella
RT haemolytica A1 glycoproteinase gene.*;
RL J. Bacteriol. 173:5597-5603(1991).
CC - FUNCTION: INVOLVED IN THE BIOSYNTHESIS OF A MOIETY ON THE CORE
CC OF THE LIPOPOLYSACCHARIDE MOLECULE.
CC - SUBCELLULAR LOCATION: CYTOPLASMIC. ASSEMBLE AT THE INNER SURFACE
CC OF THE CYTOPLASMIC MEMBRANE (POTENTIAL).
CC - SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 25.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL: U15958; AAB0283.1; -
DR InterPro: IPR002654; GL_25.
DR Pfam: PF01755; Glyco_transf_25; 1.
KW Lipopolysaccharide biosynthesis; Transferase; Glycosyltransferase.
SQ SEQUENCE 263 AA; 30871 MW; B07C39EC1DA603A7 CRC64;

Query Match 22.0%; Score 321; DB 1; Length 263;
Best Local Similarity 32.4%; Pred. No. 1.9e-22;
Matches 93; Conservative 47; Mismatches 115; Indels 32; Gaps 9;

OY 1 MONHVISLASAERRAHIAATFGSRGIPFOFDALMPSERLERAMALVGLSAHPYLSG 60
DB 1 MNHYVLSLQGRERKRIIEAFKQNPFOFQELPFDLKEKAKFNID--ISNTWLTK 58
OY 61 VERACFMSHVAWQALDEGVPIYAVFEDVLLSGAEQGLAEDTWLOERFDPDSAVVR 120
DB 59 GETACALSHLAWHLAQOMLDYICIPEDDYLGNNAFEL-----LKVYIEMNHIVK 112
OY 121 LETMFHVLTPSPGVADY--GGRAPFLISEHCGTAGYIISRKARFELDRFVLPPEL 178
DB 113 LETLPEDRINRNFKTEYILNRLFK-LNSRHGTAGIILTNKAQELIINIILTN---- 167
OY 179 HYPDLMFEGNPDREGMPYQALPALCAQELHYAFHQDSALGSLIEHRLN--RLNR-- 233
DB 168 IPIDDLFDEYELKIKKRYLQMSPALCVQDF---LINSKTNFKSSLDQDRAIRCNNEK 223
OY 234 -KOOWRDSFANTFKRLIRALTKIGREKRRQREOLIGKIIVPQ 279
DB 224 IKNOAKLKRPNYF-----LTKIQKELVPRFOLKQKVPFTYITFK 263

RESULT 9

Y765_HAEIN
ID Y765_HAEIN STANDARD; PRT; 282 AA.

AC 057125; 005033; 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical glycosyl transferase H10765 (EC 2.-.-.-).
GN H10765.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,

RA Weidmann J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uettersack T.R., Hanna M.C., Nguyen D.T., Saudak D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Georgagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Frazer C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL science 269;496-512(1995).
CC -1 SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 25.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the Embl outpostation -
CC the European Bioinformatics Institute. There are no restrictions on ways
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-stb.ch).

CC EMBL; U32760; AAC22423.1; ..
DR TIGR; H10765; ..
DR InterPro; IPR002654; GT_25.
DR Pfam; PF01755; Glyco_transf_25; 1.
DR Hypothetical protein; Transferase; Glycosyltransferase;
KW Complete proteome.
SQ SEQUENCE 282 AA; 32481 MW; D962E402AA207A18 CRC64;

Query Match 18.5%; Score 271; DB 1; Length 282;
Best Local Similarity 29.0%; Pred. No. 8.8e-18;
Matches 80; Conservative 43; Mismatches 103; Indels 50; Gaps 9

QY 3 NHVLSIAAEARRHAIATFGSRGIPFOFDDALMPSERLERMAELVPGLSNHP--YLSG 60
DB :|||||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
QY 34 NYVISLTETKORRKRIITEFEFGKNITPFFEFDAITP-DIIEFAKKFNITLIDRSPPAKISLD 92
DB :|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
QY 61 VEKACFSMSHAVMDEGALDEGVYIVAFEDDVLLSGAEBOFLADFWLGRFPDSDAFYVR 120
DB :|+|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
QY 93 GEIGCALSHVIYMDELANNNTINIFPEDDIHLGNAKELEFD-----YISDOIHYLK 146
DB :|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
QY 121 LETMFMYHTVSPGSVDYGGRAFPLLESSEHCGTGYIISRKMAREFLDFAVLPPEKLMP 180
DB :|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
DB 147 LEANGKMEFKOPKSY-KCDRRNYEPVTWRQ-SGCAGTYATAKAKEYLELVKRKPDP--VA 202
DB :|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
QY 181 VDLMFMFGPDREGNPGVOLNPALCAOE--LIYAFAFHQNNSALSLIEHDRLINKQQQR 238
DB :|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
DB 203 VDSLFEEFTLHKDKTKYTOLSPGICVOVFHH----- 234
QY 239 DSPANTEKHRLIRLTRIGREKERKRQRREDLIGKI 274
DB :|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
DB 235 --PNNPFESSL-----QEGRDRVHGNGKSSILEXI 263
DB :|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:

RESULT 10
LSG4_HAETIN STANDARD; PRT; 257 AA.

ID LSG4_HAETIN STANDARD; PRT; 257 AA.
AC P71398; Q48213;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE LSG locus putative protein 4.
DN H11697.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxId=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A2;
RC McLaughlin R., Abu Kwai Y., Young R., Spicola S., Apicella M.;
RT Characterization and sequence of the lsg locus from Haemophilus
RT Influenzae *;
RT Submitted (JUN-1992) to the EMBL/GenBank/DDBJ databases.
LN [2]
PP SEQUENCE FROM N.A.

```

RC STRAIN-Rd / RW20 / ATCC 51907;
RX MEDLINE-95350630; PubMed-7542800;
RA Flatischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kertavagis A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodet A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uutterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrman J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL collaboration
CC the European Bioinformatics Institute. There are no restrictions on its use by
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32842; AAC23343.1; -.
DR EMBL: M94855; AAA24981.1; -.
DR TIGR: H11697; -.
DR InterPro: IPR002654; GT-25.
DR Pfam: PF01755; Glyco_transf_25; 1.
DR Complete proteome.
FT FT CONFLICT 17 17 E -> K (IN REF. 1).
FT FT CONFLICT 30 30 V -> I (IN REF. 1).
FT FT CONFLICT 46 46 A -> S (IN REF. 1).
FT FT CONFLICT 102 102 S -> L (IN REF. 1).
FT FT CONFLICT 127 127 K -> N (IN REF. 1).
FT FT CONFLICT 130 130 D -> N (IN REF. 1).
FT FT CONFLICT 134 134 F -> T (IN REF. 1).
FT FT CONFLICT 153 153 D -> N (IN REF. 1).
FT FT CONFLICT 214 214 I -> M (IN REF. 1).
SQ SEQUENCE 257 AA; 29864 MW; D1ECDE47763EE059 CRC64;

Query Match. 6.9%; Score 100.5; DB 1; Length 257;
Best Local Similarity 23.0%; Pred. No. 0.046;
Matches 65; Conservative 42; Mismatches 113; Indels 63; Gaps 15;

QY 1 MONTHSLASAERAHAIAATFGSGKI-PFOFADL-MPSERLRAMAEIVPGLSAPHY 57
DB 2 LKKYILISLKDQIQRR--ELFESQKNTEDFOYFSAINIMQKMWDLALAFNIEGKAH-Y 57
QY 58 LSGVKK---ACGMSRAVLMQCALDGC---VPYIAVFEDDVLGSCADQFLAEDTWLOER 110
DB 58 FRNVYKGEIGCTLSHLYSOKIYEDNDIAEDSYALVCEDDALFHSDFOONLT--ALLSEK 115
QY 111 FDP-----DSAFVYRLQETPMFMHYLTSPSGVADVGRAFLPSEHSGTAGY 156
DB 116 LEAETIIILGQSKINDPNDPDEINPTTF-SFLCKKTKGVNV--AFP-YKSYFAGTVGY 170
QY 157 IISRKAMRFFLDREAVLPEPERLHPYDLAMFGNPDREGAPYCOLNPALCAQELAHAKFHD 216
DB 171 LIKKAARFRFQIQISQNNPFWLAD-DLFLFED-----NNNI 205
QY 217 QNSALGSL--IEHDRRLNRKQOQRDSPANTEKHLIR-ALIKYI 256
DB 206 RNKVVPRPLIVENPVLISNLESVSGSLSNMLIKIKMKYPLKRI 248

RESULT 11
EFG_CHLPN STANDARD; PRT; 694 AA.
AC Q92802; Q9JSE0;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor G (EF-G).
```

GN FUSA OR CPN0550 OR CP0202.
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia. [1]
 RX MEDLINE-99206606; PubMed-10192388;
 RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
 RA Olinger L., Gilwood J., Davis R.W., Stephens R.S.;
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.",
 RT Nat. Genet. 21:385-389(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CWL029.
 RA MEDLINE-20150255; PubMed-10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Uitterback T., Berry K., Bass S.,
 RA Linher K., Weisman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
 RT pneumoniae AR39.";
 RT Nucleic Acids Res. 28:1397-1406(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-J138;
 RA MEDLINE-20330349; PubMed-10871362;
 RA Shihai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kuhnara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 RT from Japan and CWL029 from USA.";
 RT Nucleic Acids Res. 28:2311-2314(2000).
 CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLOCATION
 CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
 CC RIBOSOME.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC EF-G/EF-2 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE001639; AAD18690.1; -
 DR EMBL: AE002181; AAF38074.1; -
 DR EMBL: AP002547; BAA98756.1; -
 DR HSSP: P13551; IDAR.
 DR PHCI-2DPAGE: 092802; -
 DR TIGR: CP0202; -
 DR InterPro: IPR004540; EF-G.
 DR InterPro: IPR000640; EF-G.
 DR InterPro: IPR004161; EFU_D2.
 DR InterPro: IPR000795; EF_GTPbind.
 DR InterPro: IPR005225; Small_GTP.
 DR Pfam: PF00009; GTP_EFTU; 1.
 DR Pfam: PF00679; EF-G; 1.
 DR PRINTS: PRO0314; GTP_EFTU_D2; 1.
 DR PRINTS: PRO0315; ELONGATNFT.
 DR TIGRFRAMS: TIGR00231; small_GTP; 1.
 DR TIGRFRAMS: TIGR00484; EF-G; 1.
 DR PROSITE: PS00301; EFATOR_GTP; 1.
 KM Elongation factor; Protein biosynthesis; GTP-binding;
 KM Complete proteome.
 FT NP_BIND 18 25 GTP (BY SIMILARITY).
 FT NP_BIND 82 86 GTP (BY SIMILARITY).
 FT NP_BIND 139 139 GTP (BY SIMILARITY).
 FT NP_BIND 555 556 NT -> KK (IN REF. 3).
 FT CONFLICT 694 AA: 76587 MW: 6AD594IDB52C87A CRC64;
 SO SEQUENCE

Query Match 6.0%; Score 88; DB 1; Length 694;
 Best Local Similarity 23.5%; Pred. No. 2.4;
 Matches 63; Conservative 33; Mismatches 78; Indels 94; Gaps 14;
 QY 41 LERMAELVPELSAPPLISGVKACFMSHAVLMDQALDEGVPTA-YFEDVLLGEGAEQ 99
 DB 93 VERSLYLDGAVAVFDVAVSGVEP-----OSETVWQADKYGVPRIAFVKMDRM--GADY 145
 QY 100 FLADTWLOERFDDSAFVVALLEMFHVLVSPSGVADYGRAPPL-----LESEHCCTA 154
 DB 146 PAAVES-WKEKL-----GANAPFVHCPIGSESQFVGWV 177
 QY 155 GYISRRAMREFLD-----REAVLPP-----ERLHPVD-----LMAF 186
 DB 178 D-LISQAL-YFLDDTIGAKWEKEISBDLKERCAELRANLELATIDESNEAFMKVYL 235
 QY 187 GNPD-----REGMPVQQLPALCAQELHYANFHDONSALGSLIEHRRILNRKQ 236
 DB 236 EDPDSITEDEIHQVWRKGVINKINPVLGTA-----KNKGVOQL-----LNVYK 282
 QY 237 WRDSPANTFKRLRLATLKIGREKER 264
 DB 283 WLSPFLDRGNIRGINLKTDEISLEPRR 310
 RESULT 12
 DDL_SYNY3
 ID DDL_SYNY3 STANDARD; PRT; 354 AA.
 AC P73632;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE D-alanine-D-alanine ligase (EC 6.3.2.4) (D-alanylanine synthetase)
 DE (D-Ala-D-Ala ligase).
 GN DDL OR DDLA OR SLR1874.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 RX NCBI_Taxid:1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97061201; PubMed-8905231;
 RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hirosewa M., Sugitara M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraiki A., Nakazaki N., Naito K.,
 RA Okumura S., Shimp S., Takeuchi C., Wada T., Watanabe A.,
 RA Yamada M., Yasuda M., Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RT DNA Res. 3:109-136(1996).
 CC -1- FUNCTION: CELL WALL FORMATION (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + D-alanine + D-alanine = ADP + phosphate
 CC + D-alanyl-D-alanine.
 CC -1- PATHWAY: ALONG WITH ALANINE RACEMASE, IT MAKES UP THE D-ALANINE
 CC BRANCH OF THE PEPTIDOGLYCAN BIOSYNTHETIC ROUTE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE D-ALANINE--D-ALANINE LIGASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D90908; BAA17677.1; -
 DR HSSP: P07862; IIOV.
 DR InterPro: IPR000291; Dala_dala_ligase.
 DR Pfam: PF01820; dala_dala_ligase; 1.
 DR TIGRFRAMS: TIGR01205; D_alan_D_alatigr; 1.
 DR PROSITE: PS00843; DALA_DALA_LIGASE_1; 1.

DR PROSITE: PS00844; DATA_LIGASE_2.1.
 KW Ligase; Cell wall; peptidoglycan synthetis; Complete proteome.
 SQ SEQUENCE 354 AA; 38774 MW; 871B83362A47315A CRC64;

Query Match 5.9%; Score 86.5; DB 1; Length 354;
 Best Local Similarity 24.3%; Pred. No. 1.4;
 Matches 43; Conservative 33; Mismatches 60; Indels 41; Gaps 11;

OY 74 EQALDEGVPIYIAVFDDVLLGEGARQFLAEDTWLQERFDPDSAFVRLTETMNVLTSPS 133
 DB 57 QCVLDGVVW-----GDQPVVTAQGRW---QFPEEA---RMEVWF-PILHGN 97
 OY 134 GYADGGRFPLLESEHCSTAGT-----ISKRAFELDRVAVLPPELHYVDL-MMFG 187
 DB 98 GGDGTVOGLFSLMQVPYVG-SCVLSCVGMCKLAKMVFER-AGLPVVMGVEKELWS 155
 OY 188 NP-----DDEGHPVCOLNPALCAQELHYAKFHDQ---SALGSLIEDRL 231
 DB 156 NCVFALCEKLEAGVGP-CVFRKANLGSSVGIKVRNRSELEALDNASYDRRI 211

RESULT 13
 PIFA_ECOLI STANDARD; PRT; 741 AA.
 ID PIFA_ECOLI
 AC P66329;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phage T7 exclusion protein.
 GN PIFA.
 OS Escherichia coli.
 OC Plasmid F.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / CR63;
 RA Shmizu H., Saitoh Y., Suda Y., Uehara K., Saempei G., Mizobuchi K.;
 RT "Complete nucleotide sequence of the F plasmid: its implications for
 organization and diversification of plasmid genomes.";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 411-735 FROM N.A.
 RC STRAIN-K12;
 RA Molineux I.J.;
 RN Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP FUNCTION.
 RA MEDLINE=71240788; PubMed=5283385;
 RA Morrison T.G., Malamy M.H.;
 RT "T7 translational control mechanisms and their inhibition by F
 factors.";
 RL Nature New Biol. 231:37-41(1971).
 -1- FUNCTION: RESPONSIBLE FOR THE EXCLUSION OF PHAGE T7 BY PLASMID F.
 GROWTH OF BACTERIOPHAGE T7 IS INHIBITED IN CELLS OF E.COLI THAT
 CARRIES THE PLASMID F.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AP001918; BAA97910.1;
 DR EMBL: U89941; AAB48033.1;
 DR Ecogen: EG40053; PIFA.
 KW Plasmid; Complete proteome.
 FT CONFLICT 701 741
 FT DOVSPFDDMPAVEKELKVMLSGNELPPEDEALKTALEND
 FT -> VTRSYSLMTGQSKRS (IN REF. 2).
 SQ SEQUENCE 741 AA; 84965 MW; 4C373DF929AFCDD5 CRC64;

Query Match 5.7%; Score 83.5; DB 1; Length 741;
 Best Local Similarity 22.1%; Pred. No. 6.9;
 Matches 75; Conservative 30; Mismatches 123; Indels 111; Gaps 16;

OY 15 RAHIAATGSRGIP-----QFDDALPSE-----RLERAAELVGLSNHPLYSGVERAC 65
 DB 184 RAELIAGQLVSLDKRIYVMDIDRLPEPSQVAEVRFLYRAVADL-----PFTTHI--LC 234
 OY 66 F-----MSHAVLMEQLDGEVPYIAVEEDVLLGEGARQFLAEDTWLQ-----EERFDDS 115
 DB 235 YDROITTHAV--EHALNT-----EDGSRLOQITIOSLFRPPEAFDLRN 277
 OY 116 AEFVRLTETMFMVLTSP--SG-----VADYGRAPFLLESEHCSTAGTYIISRAMEFF 166
 DB 278 EFRQRAEALYQOINNPPDSGKVRDLIAVTDTYGAALSTPRRIHQAINSLIPLYGMDF 337
 OY 167 LDRFVLPPERLHPVDLMFGNP-----DDEGHPV 197
 DB 338 -----VYEPD-LCLQLLRVTNPALYDWTENHYLTERSVYETGGMISDGERADFREGLIR 391
 OY 198 COLN-----PALCAQELHYAK-FHDONSALGSLIEHDRRLRKQQR-- 238
 DB 392 CMKTRASNADSEFLIADWIPGISGHNDEYLTFFPVSEDFRHIOITGKRSLSTHWRYY 451
 OY 239 ---DSPANTFKRLIRALTIGRERERKRRORREOLIGKI 274
 DB 452 FAFSPQNVLPPEFPRQLFEQAGVSEKQOQJSELLSKI 490

RESULT 14
 KBFL_MOUSE STANDARD; PRT; 971 AA.
 ID KBFL_MOUSE
 AC P25799;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Nuclear factor NF-kappa-B p105 subunit (DN-binding factor KBf1) (EBP-
 DE 1) (NF-kappa-B1 p84/NF-kappa-B1 p98) [Contains: Nuclear factor NF-
 DE kappa-B p50 subunit].
 GN NFkB1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90367113; PubMed=2203532;
 RA Chosh S., Gifford A.M., Riviere L.R., Tempet P., Nolan G.P.,
 RA Baltimore D.;
 RT "Cloning of the p50 DNA binding subunit of NF-kappa B: homology to
 RT rel and dorsal.";
 RL Cell 62:1019-1029(1990).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS P84 AND P98).
 RA MEDLINE=95059075; PubMed=7969179;
 RA Altomonte R.J., Fecondo J., Gerondakis S.;
 RT "Alternate RNA splicing of murine nfkb1 generates a nuclear isoform
 RT of the p50 precursor NF-kappa B1 that can function as a
 RT transactivator of NF-kappa B-regulated transcription.";
 RL Mol. Cell. Biol. 14:8460-8470(1994).
 RN [3]
 RP SEQUENCE OF 365-971 FROM N.A., AND PARTIAL SEQUENCE.
 RA MEDLINE=94001687; PubMed=8398903;
 RA Gerondakis S., Morrice N., Richardson I.B., Wettenhall R., Fecondo J.,
 RA Grumot R.J.;
 RT "The activity of a 70 kilodalton I kappa B molecule identical to the
 RT carboxyl terminus of the p105 NF-kappa B precursor is modulated by
 RT protein kinase A.";
 RL Cell Growth Differ. 4:617-627(1993).
 RN [4]
 RP SEQUENCE OF 365-971 FROM N.A.
 RA MEDLINE=92191289; PubMed=1339305;

DB 745 NFEELYDLDOS--WEKAGDEGVPGTTPDLMANMOVFILJNGKPYEPFVTSDDLPG 802
 OY 97 AEOFLAEDTWTLO---ERFDDSAFVYRLTFPMHVLTS 131
 DB 803 DMKQTEDETRLDCKLEIFPPDKNMATLAKRGLGIILNN 842

RESULT 15
 RET_HUMAN STANDARD; PRT; 1114 AA.
 ID P07949;
 AC 01-AUG-1988 (Rel. 08, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proto-oncogene tyrosine-protein kinase receptor ret precursor
 DE (EC 2.7.1.112) (C-ret).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE OF 1-280 FROM N.A.
 RX MEDLINE=69282215; PubMed=2660074;
 RA Takahashi M.;
 RT "Isolation of ret proto-oncogene cDNA with an amino-terminal signal
 RT sequence.";
 RL Oncogene 4:805-806(1989).
 RN [2]
 RP SEQUENCE OF 255-1114 FROM N.A.
 RX MEDLINE=90272230; PubMed=3078962;
 RA Takahashi M., Buma Y., Iwamoto T., Inaguma Y., Ikeda H., Hiai H.;
 RT "Cloning and expression of the ret proto-oncogene encoding a tyrosine
 RT kinase with two potential transmembrane domains.";
 RL Oncogene 3:571-578(1988).
 RN [3]
 RP SEQUENCE OF 588-1063 FROM N.A.
 RX MEDLINE=87257826; PubMed=3037315;
 RA Takahashi M., Cooper G.M.;
 RT "ret transforming gene encodes a fusion protein homologous to
 RT tyrosine kinases.";
 RL Mol. Cell. Biol. 7:1378-1385(1987).
 RN [4]
 RP REVIEW ON HSCR VARIANTS.
 RX MEDLINE=96023959; PubMed=9359036;
 RA Hofstra R.M.W., Oshima J., Buys C.H.C.M.;
 RT "Mutations in Hirschsprung disease: when does a mutation contribute to
 RT the phenotype.";
 RL Eur. J. Hum. Genet. 5:180-185(1997).
 RN [5]
 RP REVIEW ON VARIANTS.
 RX MEDLINE=97220587; PubMed=9067749;
 RA Eng C., Mulligan L.M.;
 RT "Mutations of the RET proto-oncogene in the multiple endocrine
 RT neoplasia type 2 syndromes, related sporadic tumours, and
 RT hirschsprung disease.";
 RL Hum. Mutat. 9:97-109(1997).
 RN [6]
 RP VARIANTS MEN2A/MTC W-611; S-618; R-620; Y-620 AND R-634.
 RX MEDLINE=93372843; PubMed=8103403;
 RA Dolits-Keller H., Dou S., Chi D., Carlson K.M., Toshima K.,
 RA Latimore T.C., Howe J.R., Moyley J.F., Goodfellow P., Wells S.A. Jr.;
 RT "Mutations in the RET proto-oncogene are associated with MEN 2A and
 RT FMTC.";
 RL Hum. Mol. Genet. 2:851-856(1993).
 RN [7]
 RP VARIANTS G-618; 632-DVR-634; G-634; F-634; Y-634 AND S-634.
 RX MEDLINE=93275414; PubMed=8099202;
 RA Mulligan L.M., Kwok J.B.J., Healey C.S., Elsdon M.J., Eng C.,
 RA Gardner E., Love D.R., Moyle S.E., Moore J.K., Papl L., Ponder M.A.,
 RA Telenius H., Tunnicliffe A., Ponder B.A.J.;
 RT "Germline mutations of the RET proto-oncogene in multiple endocrine

RT neoplasia type 2A.";
 RL Nature 363:458-460(1993).
 RN [8]
 RP VARIANTS HSCR P-40; L-399; O-762; P-765; Q-897; G-972 AND L-973.
 RX MEDLINE=95219414; PubMed=7704557;
 RA Yin L., Barone V., Seri M., Bollino A., Bocciardi R., Ceccherini I.,
 RA Pasini B., Tocco T., Lerone M., Cywes S., Moore S.,
 RA Vanderwinden J.-M., Abramowicz M.J., Kristoffersson U., Larsson L.T.,
 RA Hamel B.C.J., Silengo M., Martucciello G., Romeo G.;
 RT "Heterogeneity and low detection rate of RET mutations in Hirschsprung
 RT disease.";
 RL Eur. J. Hum. Genet. 2:272-280(1994).
 RN [9]
 RP VARIANT MEN2B THR-918.
 RX MEDLINE=94272459; PubMed=7911977;
 RA Eng C., Smith D.P., Mulligan L.M., Nagai M.A., Healey C.S.,
 RA Ponder M.A., Gardner E., Scheumann G.F., Jackson C.E., Tunnicliffe A.,
 RA Ponder B.A.J.;
 RT "Point mutation within the tyrosine kinase domain of the RET
 RT proto-oncogene in multiple endocrine neoplasia type 2B and related
 RT sporadic tumours.";
 RL Hum. Mol. Genet. 3:237-241(1994).
 RN [10]
 RP VARIANTS FMTC/MEN2A R-618; S-618; F-620; R-620; F-634; G-634 AND
 RP Y-634.
 RX MEDLINE=94348513; PubMed=7915165;
 RA Xue F., Yu H., Maurer L.H., Memoli V.A., Nuttle-Mcmenemy N.,
 RA Schuster M.K., Browden D.W., Mao J.-I., Noll M.W.;
 RT "Germline RET mutations in MEN 2A and FMTC and their detection by
 RT simple DNA diagnostic tests.";
 RL Hum. Mol. Genet. 3:635-638(1994).
 RN [11]
 RP VARIANTS FMTC/MEN2A TYR-609; ARG-618; SER-618 AND SER-620.
 RX MEDLINE=95152521; PubMed=7849720;
 RA Blaugrund J.E., Johns M.M. Jr., Eby Y.J., Ball D.W., Baylín S.B.,
 RA Hruban R.H., Sidransky D.;
 RT "RET proto-oncogene mutations in inherited and sporadic medullary
 RT thyroid cancer.";
 RL Hum. Mol. Genet. 3:1895-1897(1994).
 RN [12]
 RP VARIANTS FMTC, AND VARIANTS MEN2A.
 RX MEDLINE=95179108; PubMed=7874109;
 RA Schuffenecker I., Billaud M., Calender A., Chambe B., Ginot N.,
 RA Calmettes C., Kodigiliani E., Lenoir G.M.;
 RT "RET proto-oncogene mutations in French MEN 2A and FMTC families.";
 RL Hum. Mol. Genet. 3:1939-1943(1994).
 RN [13]
 RP VARIANT HSCR W-609, VARIANT HSCR/MEN2A R-618, AND VARIANT HSCR/FMTC
 RP R-620.
 RX MEDLINE=95187155; PubMed=7881414;
 RA Mulligan L.M., Eng C., Attie T., Lyonnet S., Marsh D.J., Hyland V.J.,
 RA Robinson B.G., Frilling A., Verellen-Dumoulin C., Sataf A.,
 RA Venter D.J., Munnich A., Ponder B.A.J.;
 RT "Diverse phenotypes associated with exon 10 mutations of the RET
 RT proto-oncogene.";
 RL Hum. Mol. Genet. 3:2163-2167(1994).
 RN [14]
 RP VARIANT MEN2B THR-918.
 RX MEDLINE=94159102; PubMed=7906866;
 RA Hofstra R.M.W., Landsvalter R.M., Ceccherini I., Stulp R.P.,
 RA Stelwagen T., Luo Y., Pasini B., Hoepfner J.W.M., van Amstel H.K.P.,
 RA Romeo G., Lips C.J.W., Buys C.H.C.M.;
 RT "A mutation in the RET proto-oncogene associated with multiple
 RT endocrine neoplasia type 2B and sporadic medullary thyroid
 RT carcinoma.";
 RL Nature 367:375-376(1994).
 RN [15]
 RP VARIANTS HSCR PRO-765; GLN-897 AND GLY-972.
 RX MEDLINE=94159103; PubMed=8114938;
 RA Romeo G., Ronchetti P., Luo Y., Barone V., Seri M., Ceccherini I.,
 RA Pasini B., Bocciardi R., Lerone M., Kaariainen H., Martucciello G.,
 RT "Point mutations affecting the tyrosine kinase domain of the RET
 RT proto-oncogene in Hirschsprung's disease.";

RL Nature 367:377-378(1994).
 RN [16]
 RP VARIANTS HSCR LEU-32; LEU-64; GLN-330 AND LEU-393.
 RX MEDLINE-94159104; PubMed-8114939;
 RA Ederly P., Lyonnet S., Mulligan L.M., Pelet A., Dow E., Abel L.,
 RT Holder S., Nihoul-Fekete C., Ponder B.A.J., Munnich A.,
 RL "Mutations of the RET proto-oncogene in Hirschsprung's disease.";
 Nature 367:378-380(1994).
 RN [17]
 RP VARIANT MEN2B THR-918.
 RX MEDLINE-94151373; PubMed-7906417;
 RA Carlson K.M., Dou S., Chi D., Scavarda N., Toshima K., Jackson C.E.,
 RT Wells S.A. Jr., Goodfellow P.J., Donis-Keller H.;
 "Single missense mutation in the tyrosine kinase catalytic domain of
 the RET protooncogene is associated with multiple endocrine neoplasia
 type 2B.";
 Proc. Natl. Acad. Sci. U.S.A. 91:1579-1583(1994).
 RN [18]
 RP VARIANTS MTC; FMTC; MEN2A AND MEN2B.
 RX MEDLINE-96223053; PubMed-8625130;
 RA Kommeth P., Kunz E.K., Matias-Guiu X., Hiort O., Christiansen G.,
 RT Colomer A., Roth J., Heitz P.U.;
 "Analysis of RET protooncogene point mutations distinguishes heritable
 from nonheritable medullary thyroid carcinomas.";
 Cancer 76:479-489(1995).
 RN [19]
 RP VARIANTS MEN2A SER-618; SER-620; ARG-634 AND TYR-634.
 RX MEDLINE-95163936; PubMed-7860065;
 RA Takiguchi-Shirahama S., Koyama K., Miyauchi A., Wakasugi T., Oishi S.,
 RT Takami H., Hiki J.K., Nakamura Y.;
 "Germline mutations of the RET proto-oncogene in eight Japanese
 patients with multiple endocrine neoplasia type 2A (MEN2A).";
 Hum. Genet. 95:187-190(1995).
 RN [20]
 RP VARIANTS HSCR L-20; S-93; Q-330; Y-609 AND R-620, AND VARIANT C-982.
 RX MEDLINE-95360000; PubMed-7633441;
 RA Angrist M., Bolk S., Thiel B., Puffenberger E.G., Hofstra R.M.W.,
 RT Buys C.H.C.M., Cass D.T., Chakravarti A.;
 "Mutation analysis of the RET receptor tyrosine kinase in Hirschsprung
 disease.";
 Hum. Mol. Genet. 4:821-830(1995).
 RN [21]
 RP VARIANTS HSCR.
 RX TISSUE-Leukocyte; PubMed-7581377;
 RA Attie T., Pelet A., Ederly P., Eng C., Mulligan L.M., Amiel J.,
 RT Boutrand L., Beldjord C., Nihoul-Fekete C., Munnich A., Ponder B.A.J.,
 RL Lyonnet S.;
 "Diversity of RET proto-oncogene mutations in familial and sporadic
 Hirschsprung disease.";
 Hum. Mol. Genet. 4:1381-1386(1995).
 RN [22]
 RP VARIANT MEN2B THR-918, AND VARIANT TYR-922.
 RX MEDLINE-96121602; PubMed-8595427;
 RA Kitamura Y., Scavarda N., Wells S.A. Jr., Jackson C.E.,
 RT Goodfellow P.J.;
 "Two maternally derived missense mutations in the tyrosine kinase
 domain of the RET protooncogene in a patient with de novo MEN 2B.";
 Hum. Mol. Genet. 4:1987-1988(1995).
 RN [23]
 RP VARIANT FMTC ASP-768.
 RX MEDLINE-95148221; PubMed-7845675;
 RA Eng C., Smith D.P., Mulligan L.M., Healey C.S., Zvelebil M.J.,
 QY 62 EKACPSHAIVLWQALDEGVPIYA---VFED---DVLLGEGAGFLADTWMQERFP 113
 DB 843 ERLTMDGLISFAMQISQGMQYLAEMKLVHRDLAARNILVAEGKMKISDFGLSRDYEE 902
 Query Match 5.7%; Score 83.5; DB 1; Length 1114;
 Best Local Similarity 19.1%; Pred. No. 12;
 Matches 41; Conservative 52; Mismatches 67; Indels 55; Gaps 10;

QY 114 DSAFVVR-----LETMFHVLITSPSGVADYGGRAFPLLESEHCGTAGIITSRKA 162
 DB 903 DS-YVRSQGRIPVKKWMAIESLFDHITYTQSDWVSFCVLLMEIV-----TLGG----- 949
 QY 163 MRFELDFANLPPERLHPVDLMMFGNPDDEGMPVQOLNPAQAEH-----YAKFHD 216
 DB 950 -----NPYEPPIPERL--FNLTKGHRMR-----PDNCSEMYRLMOCWQEPD 993
 QY 217 QNSALGSLIEH--DRRLNRKQWRDSPAHPFKHRLI 250
 DB 994 KRFVFADISKDLERKMKVKKRDYIDLAASTPSDSL 1028

Search completed: May 27, 2003, 18:14:20
 Job time : 14 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 18:12:24 ; Search time 87 Seconds
(without alignments)
660.772 Million cell updates/sec

Title: US-10-007-267a-8

Perfect score: 1461
Sequence: 1 MONHVISLASAERRAHIA.....REKRRQREQLIGKIIVPQ 279

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1408	96.4	279	2	Q93PR9	Q93PR9 neisseria g
2	1305	89.3	248	2	Q9RCN1	Q9RCN1 neisseria s
3	1284	87.9	275	2	Q9EVD4	Q9EVD4 neisseria s
4	1241	84.9	275	2	Q93EK5	Q93EK5 neisseria m
5	1111.5	76.1	280	2	Q93PR7	Q93PR7 neisseria g
6	1040.5	71.2	280	2	Q9REX4	Q9REX4 neisseria g
7	1002.5	68.6	280	2	Q9RCM8	Q9RCM8 neisseria g
8	997.5	68.3	268	16	Q9JW65	Q9JW65 neisseria m
9	967.5	66.2	280	2	P96947	P96947 neisseria m
10	931.5	63.8	280	2	Q9EVD3	Q9EVD3 neisseria s
11	912.5	62.5	280	2	Q93EK4	Q93EK4 neisseria m
12	471	32.2	266	2	Q8RKU0	Q8RKU0 haemophilus
13	439	30.0	280	2	Q917A4	Q917A4 haemophilus
14	387	26.5	282	2	Q9ZFP8	Q9ZFP8 haemophilus
15	349	23.9	208	16	Q9CLR8	Q9CLR8 pasteurella
16	252	17.2	117	2	P96944	P96944 neisseria m

17	252	17.2	133	2	Q93EK8	Q93EK8 neisseria m
18	237.5	16.3	265	16	Q98C26	Q98C26 rhizobium 1
19	231.5	15.8	268	16	Q8YRM5	Q8YRM5 brucella me
20	140.5	9.6	242	2	Q8RJX5	Q8RJX5 vibrio chol
21	131.5	9.0	257	2	Q9ZIS2	Q9ZIS2 escherichia
22	129.5	8.9	560	4	Q9P226	Q9P226 homo sapien
23	128.5	8.8	517	4	Q9DWM5	Q9DWM5 homo sapien
24	128	8.8	332	16	Q9ZLL7	Q9ZLL7 helicobacte
25	127	8.7	230	2	Q8RIJ2	Q8RIJ2 vibrio chol
26	121	8.3	1128	5	Q9VM06	Q9VM06 drosophila
27	120	8.2	207	2	Q8RJM1	Q8RJM1 vibrio chol
28	116.5	8.0	721	16	Q8TG36	Q8TG36 brucella me
29	111.5	7.6	703	2	Q8VP06	Q8VP06 brucella ab
30	110.5	7.6	287	2	Q06023	Q06023 haemophilus
31	108.5	7.4	251	2	Q9ALY2	Q9ALY2 campylobact
32	107	7.3	738	4	Q60327	Q60327 homo sapien
33	106.5	7.3	243	4	Q9H6F1	Q9H6F1 homo sapien
34	105	7.2	261	2	Q937Y0	Q937Y0 edwardsiell
35	103	7.0	224	2	Q9X605	Q9X605 actinobacil
36	102.5	7.0	231	4	Q8WU19	Q8WU19 homo sapien
37	102.5	7.0	261	16	Q9CNC3	Q9CNC3 pasteurella
38	102	7.0	254	2	Q9ALZ3	Q9ALZ3 campylobact
39	99.5	6.8	273	16	Q25500	Q25500 helicobacte
40	99	6.8	432	5	Q18989	Q18989 caenorhabdi
41	98.5	6.7	224	2	Q47960	Q47960 haemophilus
42	98.5	6.7	273	2	Q9RHG8	Q9RHG8 helicobacte
43	97	6.6	476	11	Q9DS33	Q9DS33 mus musculu
44	96	6.6	247	2	Q48023	Q48023 haemophilus
45	96	6.6	744	17	Q8ZVGI	Q8ZVGI pyrobaculum

ALIGNMENTS

RESULT 1	ID	Q93PR9	PRELIMINARY	PRT	279 AA.
AC	Q93PR9	01-DEC-2001 (TREMURel. 19, Created)			
DT	01-DEC-2001 (TREMURel. 19, Last sequence update)				
DT	01-MAR-2002 (TREMURel. 20, Last annotation update)				
DE	Glycosyl transferase LgtB.				
GN	LgtB.				
OS	Neisseria gonorrhoeae.				
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.				
OX	NCBI_TaxID=485;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
KC	STRAIN=PID2;				
RA	Tong Y., Atkling D., Reinhold V., Stein D.C.;				
RT	"Characterization of lipooligosaccharide structures found in Neisseria gonorrhoeae PID2."				
RL	Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.				
DR	EMBL: AF13394; AK70339.1; -				
DR	InterPro: IPR002654; GT_25.				
DR	Pfam: PF01755; Glyco_transf_25; 1.				
KW	Transferrase.				
SQ	SEQUENCE 279 AA; 31637 MW; 7297FL3AFA732379 CRC64;				
Query Match	96.4%; Score 1408; DB 2; Length 279;				
Best Local Similarity	96.4%; Pred. No. 8-121;				
Matches 269; Conservative	5; Mismatches 5; Indels 0; Gaps 0;				
QY	1	MONHVISLASAERRAHIAATFGSRGIPQFFDMLPSESRERAMAEVPGLSAHPYLSG 60			
DB	1	MONHVISLASAERRAHIAATFGSRGIPQFFDMLPSESRERAMAEVPGLSAHPYLSG 60			
QY	61	VEKACFSHAVLWMOALDEGVYTAIVEDVTLGEGAEQFLAEDTWLQERPDPSAVVR 120			
DB	61	VEKACFSHAVLWMOALDEGVYTAIVEDVTLGEGAEQFLAEDTWLQERPDPSAVVR 120			
QY	121	LETFEMVILVSPSGVAVYGGRAFLPILSEHCCTNGYLIISKAMFFLDRAVLPPEKLRP 180			

Db 121 LETMFMHVLTPSGVADYCGRAFPLESEHCTAGCTIISREAMPFLDFAVLPAERIK 180
 QY 181 VDLMEFGNPDDEGMPVCOALPALCAQELHYAKFHDONSALGSLIEHRLNRKQWDS 240
 Db 181 VDLMEFGNPDDEGMPVCOALPALCAQELHYAKFHDONSALGSLIEHRLNRKQWDS 240
 QY 241 PANTFKHRLIRALTKIGRERERKRRORREOLIGKIIVPQ 279
 Db 241 PANTFKHRLIRALTKIGRERERKRRORREOLIGKIIVPQ 279

RESULT 2

Q9RGNI PRELIMINARY: PRT: 248 AA.
 ID 09RGNI
 AC 09RGNI
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Glycosyltransferase.
 GN LGTB.
 OS Neisseria gonorrhoeae.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=485;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1291;
 RX MEDLINE=20305049; PubMed=10844691;
 RA Harvey H.A., Porat N., Campbell C.A., Jennings M., Gibson B.W.,
 RA Phillips N.J., Apicella M.A., Balke M.S.;
 RT "Gonococcal lipooligosaccharide is a ligand for the asialoglycoprotein
 receptor on human sperm."
 RL MOL. Microbiol. 36:1059-1070(2000).
 DR EMBL; AF121135; AAF14360.1; -.
 DR InterPro; IPR002654; GT_25.
 DR Pfam; PF01755; Glyco_transf_25; 1.
 DR Transferrase.
 KW SEQUENCE 248 AA; 28036 MW; D9036FB07D8C9397 CRC64;

Query Match 89.3%; Score 1305; DB 2; Length 248;
 Best Local Similarity 99.6%; Pred. No. 2e-111;
 Matches 247; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MONHVISLASAERRAHIAATFGSGRGIPOFFDALMPSERLERAMAEIVPGISAHPTLSG 60
 Db 1 MONHVISLASAERRAHIAATFGSGRGIPOFFDALMPSERLERAMAEIVPGISAHPTLSG 60
 QY 61 VEKACFMSHAYLMQALDEGVPIYAVFEDDVLLEGAGQFLAEDTWLOERPDPSAFYVR 120
 Db 61 VEKACFMSHAYLMQALDEGVPIYAVFEDDVLLEGAGQFLAEDTWLOERPDPSAFYVR 120
 QY 121 LETMFMHVLTPSGVADYCGRAFPLESEHCTAGCTIISRKAMRFFLDRAVLPPERLHP 180
 Db 121 LETMFMHVLTPSGVADYCGRAFPLESEHCTAGCTIISRKAMRFFLDRAVLPPERLHP 180
 QY 181 VDLMEFGNPDDEGMPVCOALPALCAQELHYAKFHDONSALGSLIEHRLNRKQWDS 240
 Db 181 VDLMEFGNPDDEGMPVCOALPALCAQELHYAKFHDONSALGSLIEHRLNRKQWDS 240
 QY 241 PANTFKHR 248
 Db 241 PANTFKHR 248

RESULT 3

Q9EVD4 PRELIMINARY: PRT: 275 AA.
 ID 09EVD4
 AC 09EVD4
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Galactosyl transferase.
 GN LGTB.
 OS Neisseria subflava.

OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=28449;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21142520; PubMed=11208792;
 RA Arking D., Tong Y., Stein D.C.;
 RT "Analysis of lipooligosaccharide biosynthesis in the Neisseriaceae."
 RL J. Bacteriol. 183:934-941(2001).
 DR EMBL; AF240672; AAG09765.1; -.
 DR InterPro; IPR002654; GT_25.
 DR Pfam; PF01755; Glyco_transf_25; 1.
 DR Transferrase.
 KW SEQUENCE 275 AA; 31441 MW; D719E3815F64D14C CRC64;

Query Match 87.9%; Score 1284; DB 2; Length 275;
 Best Local Similarity 86.9%; Pred. No. 1.9e-109;
 Matches 248; Conservative 9; Mismatches 18; Indels 4; Gaps 1;

QY 1 MONHVISLASAERRAHIAATFGSGRGIPOFFDALMPSERLERAMAEIVPGISAHPTLSG 60
 Db 1 MONHVISLASAERRAHIAATFGSGRGIPOFFDALMPSERLERAMAEIVPGISAHPTLSG 60
 QY 61 VEKACFMSHAYLMQALDEGVPIYAVFEDDVLLEGAGQFLAEDTWLOERPDPSAFYVR 120
 Db 61 VEKACFMSHAYLMQALDEGVPIYAVFEDDVLLEGAGQFLAEDTWLOERPDPSAFYVR 120
 QY 121 LETMFMHVLTPSGVADYCGRAFPLESEHCTAGCTIISRKAMRFFLDRAVLPPERLHP 180
 Db 121 LETMFMHVLTPSGVADYCGRAFPLESEHCTAGCTIISRKAMRFFLDRAVLPPERLHP 180
 QY 181 VDLMEFGNPDDEGMPVCOALPALCAQELHYAKFHDONSALGSLIEHRLNRKQWDS 240
 Db 181 VDLMEFGNPDDEGMPVCOALPALCAQELHYAKFHDONSALGSLIEHRLNRKQWDS 240
 QY 241 PANTFKHRLIRALTKIGRERERKRRORREOLIGKIIVPQ 279
 Db 241 PANTFKHRLIRALTKIGRERERKRRORREOLIGKIIVPQ 279

RESULT 4

Q93EK5 PRELIMINARY: PRT: 275 AA.
 ID 093EK5
 AC 093EK5
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE LGTB.
 GN LGTB.
 OS Neisseria meningitidis.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=487;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M978;
 RX MEDLINE=21467954; PubMed=11583844;
 RA Zhu P., Kutuch M.J., Tsai C.-M.;
 RT "Genetic Analysis of Conservation and Variation of Lipooligosaccharide
 Expression in Two L8-Immunotype Strains of Neisseria meningitidis."
 RL FEMS Microbiol. Lett. 203:173-177(2001).
 DR EMBL; AF355193; AAL12841.1; -.
 DR InterPro; IPR002654; GT_25.
 DR Pfam; PF01755; Glyco_transf_25; 1.
 DR Transferrase.
 KW SEQUENCE 275 AA; 31577 MW; A73037E0625EE7B3 CRC64;

Query Match 84.9%; Score 1241; DB 2; Length 275;
 Best Local Similarity 86.4%; Pred. No. 1.6e-105;
 Matches 241; Conservative 17; Mismatches 17; Indels 4; Gaps 1;

QY 1 MONHVISLASAERRAHIAATFGSGRGIPOFFDALMPSERLERAMAEIVPGISAHPTLSG 60
 Db 1 MONHVISLASAERRAHIAATFGSGRGIPOFFDALMPSERLERAMAEIVPGISAHPTLSG 60
 QY 61 VEKACFMSHAYLMQALDEGVPIYAVFEDDVLLEGAGQFLAEDTWLOERPDPSAFYVR 120

```

DB 61 VERACFMSHAVLWEOALDEGLPIAVFEDVLLGEGAEKFLADDAWLKEFPDPSAFYR 120
      |||
OY 121 LETMFHVLTPSGVADYGGRAFPLESEHCAGTYIISRKARFPLDRVAVPPERLHP 180
      |||
DB 121 LETMFHVLTPSGVADYGGRAFPLESEHCAGTYIISRKARFPLDRVAVPPERLHP 180
      |||
OY 181 VDLMEFGNPDREGMPVQCOLNPALCAOELHYAFHONSALGSLIEHRLNRKQQRDS 240
      |||
DB 181 VDLMEFGNPDREGMPVQCOLNPALCAOELHYAFHONSALGSLIEHRLNRKQQRDS 240
      |||
OY 241 PANTFMRRLRLATKTRERERKRRORREOLIGTIYFQ 279
      |||
DB 241 PANTFMRRLRLATKTRERERKRRORREOLIGTIYFQ 279
      |||

RESULT 5
O93PR7 PRELIMINARY; PRT; 280 AA.
AC O93PR7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Glycosyl transferase LGTE.
GN LGTE.
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PID2;
RA Tong Y., Atkling D., Reinhold V., Stein D.C.;
RT "Characterization of lipooligosaccharide structures found in Neisseria
   gonorrhoeae PID2."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF313394; AKK70340.1; -
DR InterPro: IPR002654; GT_25.
DR Pfam: PF01755; Glyco_transf_25; 1.
DR Transferase.
KW SEQUENCE 280 AA; 32047 MW; 4F8171AC1BACFB45 CRC64;

Query Match 76.1%; Score 1111.5; DB 2; Length 280;
Best Local Similarity 76.7%; Pred. No. 1.2e-93;
Matches 224; Conservative 19; Mismatches 24; Indels 25; Gaps 3;

OY 1 MONHVISLASAERRAHIAATFGSRGIPFOFPALMPSERLERAMAEVGLSAHPYLSG 60
      |||
DB 1 MONHVISLASAERRAHIAATFGSRGIPFOFPALMPSERLERAMAEVGLSAHPYLSG 60
      |||
OY 61 VERACFMSHAVLWEOALDEGVPIAVFEDVLLGEGAEQFLAEDTWLQERFPDPSAFVR 120
      |||
DB 61 VERACFMSHAVLWEOALDEGVPIAVFEDVLLGEGAEQFLAEDTWLQERFPDPSAFVR 120
      |||
OY 121 LETMFHVLTPSGVADYGGRAFPLESEHCAGTYIISRKARFPLDRVAVPPERLHP 180
      |||
DB 121 LETMFHVLTPSGVADYGGRAFPLESEHCAGTYIISRKARFPLDRVAVPPERLHP 180
      |||
OY 181 VDLMEFGNPDREGMPVQCOLNPALCAOELHYAFHONSALGSLIEHRLNRKQQRDS 240
      |||
DB 181 VDLMEFGNPDREGMPVQCOLNPALCAOELHYAFHONSALGSLIEHRLNRKQQRDS 240
      |||
OY 241 PANTFMRRLRLATKTRERERKRRORREOLIGTIYFQ 279
      |||
DB 241 PANTFMRRLRLATKTRERERKRRORREOLIGTIYFQ 279
      |||

RESULT 6
O9REX4 PRELIMINARY; PRT; 280 AA.
AC O9REX4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

```

```

DE LGTE.
GN LGTE.
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FA19;
RA Baltazar J.T., Shafer W.M., Stephens D.S., Martin L.E.;
RT "Mutations in the lgt operon influence serum-resistance in
   gonococci."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF208063; AAF20992.1; -
DR InterPro: IPR002654; GT_25.
DR Pfam: PF01755; Glyco_transf_25; 1.
DR Transferase.
KW SEQUENCE 280 AA; 32448 MW; A20B34CD786942E3 CRC64;

Query Match 71.2%; Score 1040.5; DB 2; Length 280;
Best Local Similarity 71.9%; Pred. No. 3.7e-87;
Matches 210; Conservative -23; Mismatches 34; Indels 25; Gaps 3;

OY 1 MONHVISLASAERRAHIAATFGSRGIPFOFPALMPSERLERAMAEVGLSAHPYLSG 60
      |||
DB 1 MONHVISLASAERRAHIAATFGSRGIPFOFPALMPSERLERAMAEVGLSAHPYLSG 60
      |||
OY 61 VERACFMSHAVLWEOALDEGVPIAVFEDVLLGEGAEQFLAEDTWLQERFPDPSAFVR 120
      |||
DB 61 VERACFMSHAVLWEOALDEGVPIAVFEDVLLGEGAEQFLAEDTWLQERFPDPSAFVR 120
      |||
OY 121 LETMFHVLTPSGVADYGGRAFPLESEHCAGTYIISRKARFPLDRVAVPPERLHP 180
      |||
DB 121 LETMFHVLTPSGVADYGGRAFPLESEHCAGTYIISRKARFPLDRVAVPPERLHP 180
      |||
OY 181 VDLMEFGNPDREGMPVQCOLNPALCAOELHYAFHONSALGSLIEHRLNRKQQRDS 240
      |||
DB 181 VDLMEFGNPDREGMPVQCOLNPALCAOELHYAFHONSALGSLIEHRLNRKQQRDS 240
      |||
OY 241 PANTFMRRLRLATKTRERERKRRORREOLIGTIYFQ 279
      |||
DB 241 PANTFMRRLRLATKTRERERKRRORREOLIGTIYFQ 279
      |||

RESULT 7
O9RGMB PRELIMINARY; PRT; 280 AA.
AC O9RGMB;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Glycosyltransferase.
GN LGTE.
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1291;
RA MEDLINE=20305049; PubMed=10844691;
RA Harvey H.A., Porat N., Campbell C.A., Jennings M., Gibson B.W.,
RA Phillips N.J., Apicella M.A., Balke M.S.;
RT "Gonococcal lipooligosaccharide is a ligand for the asialoglycoprotein
   receptor on human sperm."
RL Mol. Microbiol. 36:1059-1070(2000).
DR EMBL: AF121135; AAF14363.1; -
DR InterPro: IPR002654; GT_25.
DR Pfam: PF01755; Glyco_transf_25; 1.
DR Transferase.
KW SEQUENCE 280 AA; 32304 MW; 2EB5CDE7D2164E6 CRC64;

Query Match 68.6%; Score 1002.5; DB 2; Length 280;
Best Local Similarity 69.9%; Pred. No. 1.1e-83;
Matches 204; Conservative 23; Mismatches 40; Indels 25; Gaps 3;

```

```

OY 1 MONHYISLASAERRAHATATGSGIPQFFDALMPSERLERAAELVPGLSAHPYLSG 60
DB 1 MONHYISLASAERRAHATADTFGSGIPQFFDALMPSERLERAAELVPGLSAHPYLSG 60
OY 61 VKACFMSHAVALMEQALDEGVYIAVFEDDVLGGABQFLAEDTWLOERPDPSAFYVR 120
DB 61 VKACFMSHAVALMEQALDEGVYIAVFEDDVLGGABQFLAEDTWLOERPDPSAFYVR 120
OY 121 LETMFMHVLTPSGVADYGGRAFPLESEHGCTAGYIISRKAMRFLDRFAVLPPERLHP 180
DB 121 LETMFMHVLTPSGVADYGGRAFPLESEHGCTAGYIISRKAMRFLDRFAVLPPERLHP 180
OY 121 LETMFMHVLTPSGVADYGGRAFPLESEHGCTAGYIISRKAMRFLDRFAVLPPERLHP 180
DB 121 LETMFMHVLTPSGVADYGGRAFPLESEHGCTAGYIISRKAMRFLDRFAVLPPERLHP 180
OY 181 VDLMEGNDDEGMPVQOLNPAALCAQELHYAKFHONASLSLIEHRRRLNRKQOWDS 240
DB 181 VDLMEGNDDEGMPVQOLNPAALCAQELHYAKFHONASLSLIEHRRRLNRKQOWDS 240
OY 241 PANTFKHR-----LIRALTKIGRERERK-----RQREOLIGKIYVPO 279
DB 241 PANTFKHR-----LIRALTKIGRERERK-----RQREOLIGKIYVPO 279
OY 241 PANTFKHR-----LIRALTKIGRERERK-----RQREOLIGKIYVPO 279
DB 241 PANTFKHR-----LIRALTKIGRERERK-----RQREOLIGKIYVPO 279

```

RESULT 8

```

OY 09JW65 PRELIMINARY; PRT; 268 AA.
AC 09JW65:
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Lacto-N-neotetraose biosynthesis glycosyl transferase (LgtH).
GN LGT2 OR NMA0527 OR LGTH.
OS Neisseria meningitidis (serogroup A), and
OS Neisseria meningitidis
OC Bacteria, Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=68699; 487;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=N.meningitidis (serogroup A);
RC STRAIN=2491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jajelski K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491."
RT Nature 404:502-506(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=N.meningitidis; STRAIN=AI;
RC MEDLINE=21467934; PubMed=11583844;
RA Zhu P., Kutuch M.J., Tsai C.-M.;
RT "Genetic Analysis of Conservation and Variation of Lipooligosaccharide
RT Expression in Two 18-Immunotype Strains of Neisseria meningitidis."
RL FEMS Microbiol. Lett. 203:173-177(2001).
DR EMBL; AL162753; CAB83819.1;
DR EMBL; AF35194; AAL12844.1;
DR InterPro; IPR002654; GT_25;
DR Pfam; PF01755; Glyco_transf_25; 1.
KW Complete proteome.
SQ SEQUENCE 268 AA; 30444 MW; 8F43967D0530DD02.CRC64;

```

Query Match 68.3%; Score 997.5; DB 16; Length 268;
 Best Local Similarity 72.9%; Pred. No. 3e-83;
 Matches 207; Conservative 24; Mismatches 32; Indels 21; Gaps 5;

```

OY 1 MONHYISLASAERRAHATATGSGIPQFFDALMPSERLERAAELVPGLSAHPYLSG 60
DB 1 MONHYISLASAERRAHATADTFGSGIPQFFDALMPSERLERAAELVPGLSAHPYLSG 60
OY 61 VKACFMSHAVALMEQALDEGVYIAVFEDDVLGGABQFLAEDTWLOERPDPSAFYVR 120
DB 61 VKACFMSHAVALMEQALDEGVYIAVFEDDVLGGABQFLAEDTWLOERPDPSAFYVR 120

```

```

DB 61 VKACFMSHAVALMEQALDEGVYIAVFEDDVLGGABQFLAEDTWLOERPDPSAFYVR 120
OY 121 LETMFMHVLTPSGVADYGGRAFPLESEHGCTAGYIISRKAMRFLDRFAVLPPERLHP 180
DB 121 LETMFMHVLTPSGVADYGGRAFPLESEHGCTAGYIISRKAMRFLDRFAVLPPERLHP 180
OY 181 VDLMEGNDDEGMPVQOLNPAALCAQELHYAKFHONASLSLIEHRRRLNRKQOWDS 240
DB 181 VDLMEGNDDEGMPVQOLNPAALCAQELHYAKFHONASLSLIEHRRRLNRKQOWDS 240
OY 241 PANTFKHR-----LIRALTKIGRERERK-----RQREOLIGKIYVPO 279
DB 241 PANTFKHR-----LIRALTKIGRERERK-----RQREOLIGKIYVPO 279
OY 241 PANTFKHR-----LIRALTKIGRERERK-----RQREOLIGKIYVPO 279
DB 241 PANTFKHR-----LIRALTKIGRERERK-----RQREOLIGKIYVPO 279

```

RESULT 9

```

OY 09EVD3 PRELIMINARY; PRT; 280 AA.
AC 09EVD3:
DT 01-MAY-1997 (TREMblrel. 03, Created)
DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE Glycosyl transferase.
GN LGTE.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=126E;
RC MEDLINE=20055626; PubMed=10589709;
RX Jennings M.P., Srikantha Y.N., Moxon E.R., Kramer M., Poolman J.T.,
RA Kuipers B., Van der Ley P.;
RT "The genetic basis of the phase variation repertoire of
RT lipopolysaccharide immunotypes in Neisseria meningitidis."
RL Microbiology 145:3013-3021(1999).
DR EMBL; U65788; AAB48387.1;
DR InterPro; IPR002654; GT_25;
DR Pfam; PF01755; Glyco_transf_25; 1.
KW Transferase.
SQ SEQUENCE 280 AA; 32589 MW; 1ED618E2E08AFCA.CRC64;

```

Query Match 66.2%; Score 967.5; DB 2; Length 280;
 Best Local Similarity 67.1%; Pred. No. 1.8e-80;
 Matches 196; Conservative 29; Mismatches 42; Indels 25; Gaps 3;

```

OY 1 MONHYISLASAERRAHATATGSGIPQFFDALMPSERLERAAELVPGLSAHPYLSG 60
DB 1 MONHYISLASAERRAHATADTFGSGIPQFFDALMPSERLERAAELVPGLSAHPYLSG 60
OY 61 VKACFMSHAVALMEQALDEGVYIAVFEDDVLGGABQFLAEDTWLOERPDPSAFYVR 120
DB 61 VKACFMSHAVALMEQALDEGVYIAVFEDDVLGGABQFLAEDTWLOERPDPSAFYVR 120
OY 121 LETMFMHVLTPSGVADYGGRAFPLESEHGCTAGYIISRKAMRFLDRFAVLPPERLHP 180
DB 121 LETMFMHVLTPSGVADYGGRAFPLESEHGCTAGYIISRKAMRFLDRFAVLPPERLHP 180
OY 121 LETMFMHVLTPSGVADYGGRAFPLESEHGCTAGYIISRKAMRFLDRFAVLPPERLHP 180
DB 121 LETMFMHVLTPSGVADYGGRAFPLESEHGCTAGYIISRKAMRFLDRFAVLPPERLHP 180
OY 181 VDLMEGNDDEGMPVQOLNPAALCAQELHYAKFHONASLSLIEHRRRLNRKQOWDS 240
DB 181 VDLMEGNDDEGMPVQOLNPAALCAQELHYAKFHONASLSLIEHRRRLNRKQOWDS 240
OY 241 PANTFKHR-----LIRALTKIGRERERK-----RQREOLIGKIYVPO 279
DB 241 PANTFKHR-----LIRALTKIGRERERK-----RQREOLIGKIYVPO 279
OY 241 PANTFKHR-----LIRALTKIGRERERK-----RQREOLIGKIYVPO 279
DB 241 PANTFKHR-----LIRALTKIGRERERK-----RQREOLIGKIYVPO 279

```

RESULT 10

```

OY 09EVD3 PRELIMINARY; PRT; 280 AA.
AC 09EVD3:
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)

```

DT 01-MAR-2002 (Tremblrel. 20, last annotation update)
 DE Galactosyl transferase.
 GN LGTE.
 OS Neisseria.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_Taxid=28449;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2142520; PubMed=11208792;
 RA Arking D., Tong Y., Stein D.C.;
 RT "Analysis of lipooligosaccharide biosynthesis in the Neisseriaceae."
 RL J. Bacteriol. 183:934-941(2001).
 DR EMBL; AF240672; AAG09766.1; -
 DR InterPro; IPR002654; GT_25.
 DR Pfam; PF01755; Glyco_transf_25; 1.
 KW Transference.
 SQ SEQUENCE 280 AA; 32762 MW; 85CF77E94F215842 CRC64;

Query Match 63.8%; Score 931.5; DB 2; Length 280;
 Best Local Similarity 65.4%; Pred. No. 3.5e-77;
 Matches 191; Conservative 29; Mismatches 47; Indels 25; Gaps 3;

QY 1 MONHVLSLSAARAHIAATFGSGRIPQFPDAMPSELERAMALVPGLSAHPYLSG 60
 DB 1 MONHVLSLSAARAHIAATFGSGRIPQFPDAMPSELERAMALVPGLSAHPYLSG 60
 QY 61 VERACFMSHAVLMEQALDEGVPIAVFEDVLLGEGAEQFLADDTWLOERFPDPSAFVVR 120
 DB 61 VERACFMSHAVLMEQALDEGVPIAVFEDVLLGEGAEQFLADDTWLOERFPDPSAFVVR 120
 QY 121 LETMFHVLTPSPGVADYCGRAFPPLLESEHCAGTAGYIISRKARFELDRFVLPPELHP 180
 DB 121 LETMFHVLTPSPGVADYCGRAFPPLLESEHCAGTAGYIISRKARFELDRFVLPPELHP 180
 QY 121 LETMFHVLTPSPGVADYCGRAFPPLLESEHCAGTAGYIISRKARFELDRFVLPPELHP 180
 DB 121 LETMFHVLTPSPGVADYCGRAFPPLLESEHCAGTAGYIISRKARFELDRFVLPPELHP 180
 QY 181 VDLAMFNGNDREGMPVQOLNPAALCAQELHYAKFHONSALGSLIEHDLNRKQWRDS 240
 DB 181 VDLAMFNGNDREGMPVQOLNPAALCAQELHYAKFHONSALGSLIEHDLNRKQWRDS 240
 QY 241 PANTFKR-----LIRALTKIGRERERORR-----EOLIGKIIVPQ 279
 DB 241 PANTFKR-----LIRALTKIGRERERORR-----EOLIGKIIVPQ 279
 QY 234 -----RHRSLSKVMFDLKRALGKFGREKKRMERQRALEKAYGRVVISFK 280
 DB 234 -----RHRSLSKVMFDLKRALGKFGREKKRMERQRALEKAYGRVVISFK 280

RESULT 11
 QY3EK4 PRELIMINARY; PRT; 280 AA.

AC 093EK4;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, last annotation update)
 DE LGTE.
 GN LGTE.
 OS Neisseria meningitidis.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_Taxid=487;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21467954; PubMed=11583844;
 RA Zhu P., Klutch M.J., Tsai C.-M.;
 RT "Genetic Analysis of Conservation and Variation of Lipooligosaccharide Expression in Two L8-ImmunoType Strains of Neisseria meningitidis."
 RL FEMS Microbiol. Lett. 203:173-177(2001).
 DR EMBL; AF355193; AAL12842.1; -
 DR InterPro; IPR002654; GT_25.
 DR Pfam; PF01755; Glyco_transf_25; 1.
 KW Transference.
 SQ SEQUENCE 280 AA; 32898 MW; 504256FF5D92F97 CRC64;

Query Match 62.5%; Score 912.5; DB 2; Length 280;
 Best Local Similarity 64.4%; Pred. No. 1.9e-75;
 Matches 188; Conservative 28; Mismatches 51; Indels 25; Gaps 3;

QY 1 MONHVLSLSAARAHIAATFGSGRIPQFPDAMPSELERAMALVPGLSAHPYLSG 60

DB 1 MONHVLSLSAARAHIAATFGSRHDIPEQFPALMPSELERAMALVPGLSAHPYLSG 60
 QY 61 VERACFMSHAVLMEQALDEGVPIAVFEDVLLGEGAEQFLADDTWLOERFPDPSAFVVR 120
 DB 61 VERACFMSHAVLMEQALDEGVPIAVFEDVLLGEGAEQFLADDTWLOERFPDPSAFVVR 120
 QY 121 LETMFHVLTPSPGVADYCGRAFPPLLESEHCAGTAGYIISRKARFELDRFVLPPELHP 180
 DB 121 LETMFHVLTPSPGVADYCGRAFPPLLESEHCAGTAGYIISRKARFELDRFVLPPELHP 180
 QY 121 LETMFHVLTPSPGVADYCGRAFPPLLESEHCAGTAGYIISRKARFELDRFVLPPELHP 180
 DB 121 LETMFHVLTPSPGVADYCGRAFPPLLESEHCAGTAGYIISRKARFELDRFVLPPELHP 180
 QY 181 VDLAMFNGNDREGMPVQOLNPAALCAQELHYAKFHONSALGSLIEHDLNRKQWRDS 240
 DB 181 VDLAMFNGNDREGMPVQOLNPAALCAQELHYAKFHONSALGSLIEHDLNRKQWRDS 240
 QY 241 PANTFKR-----LIRALTKIGRERERORR-----EOLIGKIIVPQ 279
 DB 241 PANTFKR-----LIRALTKIGRERERORR-----EOLIGKIIVPQ 279
 QY 234 -----RHRSLSKVMFDLKRALGKFGREKKRMERQRALEKAYGRVVISFK 280
 DB 234 -----RHRSLSKVMFDLKRALGKFGREKKRMERQRALEKAYGRVVISFK 280

RESULT 12
 QY3EK4 PRELIMINARY; PRT; 266 AA.

AC 08RKU0;
 DT 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
 DE Putative galactosyl transferase Lfc2B.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_Taxid=727;
 RN
 RP SEQUENCE FROM N.A.
 RX STRAIN=6622;
 RA Pettigrew M.M., Foxman B., Mairs C.F., Gilsdorf J.R.;
 RT "Identification of LOS biosynthesis gene lfc2B as a putative virulence factor in strains of non-typable Haemophilus influenzae that cause otitis media."
 RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY091470; AAM12037.1; -
 KW Transference.
 SQ SEQUENCE 266 AA; 31182 MW; B2ICPA372200E2DC CRC64;

Query Match 32.2%; Score 471; DB 2; Length 266;
 Best Local Similarity 39.8%; Pred. No. 4.4e-35;
 Matches 111; Conservative 46; Mismatches 104; Indels 18; Gaps 6;

QY 3 NHVLSLSAARAHIAATFGSGRIPQFPDAMPSELERAMALVPGLSAHPYLSGVE 62
 DB 3 NHVLSLSAARAHIAATFGSGRIPQFPDAMPSELERAMALVPGLSAHPYLSGVE 62
 QY 4 NYVLSLSAARAHIAATFGSGRIPQFPDAMPSELERAMALVPGLSAHPYLSGVE 62
 DB 4 NYVLSLSAARAHIAATFGSGRIPQFPDAMPSELERAMALVPGLSAHPYLSGVE 62
 QY 63 KACFMSHAVLMEQALDEGVPIAVFEDVLLGEGAEQFLADDTWLOERFPDPSAFVVR 122
 DB 63 KACFMSHAVLMEQALDEGVPIAVFEDVLLGEGAEQFLADDTWLOERFPDPSAFVVR 122
 QY 123 TMEFHVLTSPGVADYCGRAFPPLLESEHCAGTAGYIISRKARFELDRFVLPPELHP 182
 DB 123 TMEFHVLTSPGVADYCGRAFPPLLESEHCAGTAGYIISRKARFELDRFVLPPELHP 182
 QY 123 TMEFHVLTSPGVADYCGRAFPPLLESEHCAGTAGYIISRKARFELDRFVLPPELHP 182
 DB 123 TMEFHVLTSPGVADYCGRAFPPLLESEHCAGTAGYIISRKARFELDRFVLPPELHP 182
 QY 183 LAMFNGNDREGMPVQOLNPAALCAQELHYAKFHONSALGSLIEHDLNRKQWRDS 240
 DB 183 LAMFNGNDREGMPVQOLNPAALCAQELHYAKFHONSALGSLIEHDLNRKQWRDS 240
 QY 241 PANTFKR-----LIRALTKIGRERERORR-----EOLIGKIIVPQ 279
 DB 241 PANTFKR-----LIRALTKIGRERERORR-----EOLIGKIIVPQ 279
 QY 232 --RTLKVTLISLAGK--PKILKRIYKLIKSHIVPFR 266
 DB 232 --RTLKVTLISLAGK--PKILKRIYKLIKSHIVPFR 266

RESULT 13
 QY3L7A4 PRELIMINARY; PRT; 280 AA.

DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Galactosyltransferase II.
 GN lgtB.
 OS Haemophilus ducreyi.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=730;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=35000HP;
 RX MEDLINE=20200369; PubMed=10735874;
 RA Sun S., Schilling B., Tarantino L., Tullius M.V., Gibson B.W.,
 Munson R.S. Jr.;
 RT "Cloning and characterization of the lipooligosaccharide
 galactosyltransferase II gene of Haemophilus ducreyi.";
 RL J. Bacteriol. 182:2292-2298(2000).
 DR EMBL; AF224466; AAF32397.1;
 DR InterPro; IPR002654; GT_25.
 DR Pfam; PF01755; Glyco_transf_25; 1.
 KM Transferase.
 SQ SEQUENCE 280 AA; 33369 MW; 0639F668C1F5E6E CRC64;
 Query Match 30.0%; Score 439; DB 2; Length 280;
 Best Local Similarity 35.7%; Pred. No. 4e-32;
 Matches 100; Conservative 59; Mismatches 111; Indels 10; Gaps 4;

OY 3 NHVTSLSAARRAHIAATFGSRGIPFDFALMPSERLERAMAEVPGLSAHPYLSGVE 62
 ID 09CLR8 PRELIMINARY; PRT; 282 AA.
 AC 09CLR8;
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Glycosyltransferase homolog lcb2A.
 GN lcb2A.
 OS Haemophilus somnus.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=731;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=738;
 RX MEDLINE=20072700; PubMed=10603403;
 RA Wu Y., McQuiston J.H., Cox A., Pack T.D., Inzana T.J.;
 RT "Molecular cloning and mutagenesis of a DNA locus involved in
 lipooligosaccharide biosynthesis in Haemophilus somnus.";
 RL Infect. Immun. 68:310-319(2000).
 DR EMBL; AF096997; AAC83803.1;
 DR InterPro; IPR002654; GT_25.
 DR Pfam; PF01755; Glyco_transf_25; 1.

OY 183 LMFGNPDRCMPVCOINPALCAOELHYAKFHQNSALGSLIEHRLNRRQWRDSPA 242
 ID 09CLR8 PRELIMINARY; PRT; 282 AA.
 AC 09CLR8;
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Glycosyltransferase homolog lcb2A.
 GN lcb2A.
 OS Haemophilus somnus.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=731;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=738;
 RX MEDLINE=20072700; PubMed=10603403;
 RA Wu Y., McQuiston J.H., Cox A., Pack T.D., Inzana T.J.;
 RT "Molecular cloning and mutagenesis of a DNA locus involved in
 lipooligosaccharide biosynthesis in Haemophilus somnus.";
 RL Infect. Immun. 68:310-319(2000).
 DR EMBL; AF096997; AAC83803.1;
 DR InterPro; IPR002654; GT_25.
 DR Pfam; PF01755; Glyco_transf_25; 1.

OY 243 NTFKRLRLA--LTGIREREKRORREQLIGKIIVPQ 279
 ID 09CLR8 PRELIMINARY; PRT; 282 AA.
 AC 09CLR8;
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Glycosyltransferase homolog lcb2A.
 GN lcb2A.
 OS Haemophilus somnus.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=731;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=738;
 RX MEDLINE=20072700; PubMed=10603403;
 RA Wu Y., McQuiston J.H., Cox A., Pack T.D., Inzana T.J.;
 RT "Molecular cloning and mutagenesis of a DNA locus involved in
 lipooligosaccharide biosynthesis in Haemophilus somnus.";
 RL Infect. Immun. 68:310-319(2000).
 DR EMBL; AF096997; AAC83803.1;
 DR InterPro; IPR002654; GT_25.
 DR Pfam; PF01755; Glyco_transf_25; 1.

KW Transferase.
 SQ SEQUENCE 282 AA; 33687 MW; 494094032C7024DD CRC64;
 Query Match 26.5%; Score 387; DB 2; Length 282;
 Best Local Similarity 35.5%; Pred. No. 2.3e-27;
 Matches 102; Conservative 47; Mismatches 120; Indels 18; Gaps 6;

OY 1 MONVTSLSAARRAHIAATFGSRGIPFDFALMPSERLERAMAEVPGLSAHPYLSG 60
 ID 09CLR8 PRELIMINARY; PRT; 208 AA.
 AC 09CLR8;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Hypothetical protein Pm141.
 GN Pm141.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Pasteurella.
 OX NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PM70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida Pm70.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 DR EMBL; AE006155; AAK03225.1;
 DR InterPro; IPR002654; GT_25.
 DR Pfam; PF01755; Glyco_transf_25; 1.
 DR Hypothetical protein; Complete proteome.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 208 AA; 24649 MW; 29625BA115E4047A CRC64;
 Query Match 23.9%; Score 349; DB 16; Length 208;
 Best Local Similarity 37.3%; Pred. No. 4.7e-24;
 Matches 82; Conservative 40; Mismatches 74; Indels 24; Gaps 6;

OY 72 LMEOALDEGVPIYAVEEDVLLGEGAEQFLAEDTWLOERPDPSAFVYRLTGMHVLTS 131
 ID 09CLR8 PRELIMINARY; PRT; 208 AA.
 AC 09CLR8;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Hypothetical protein Pm141.
 GN Pm141.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Pasteurella.
 OX NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PM70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida Pm70.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 DR EMBL; AE006155; AAK03225.1;
 DR InterPro; IPR002654; GT_25.
 DR Pfam; PF01755; Glyco_transf_25; 1.
 DR Hypothetical protein; Complete proteome.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 208 AA; 24649 MW; 29625BA115E4047A CRC64;
 Query Match 23.9%; Score 349; DB 16; Length 208;
 Best Local Similarity 37.3%; Pred. No. 4.7e-24;
 Matches 82; Conservative 40; Mismatches 74; Indels 24; Gaps 6;

OY 132 --PSCVADYGGRAFPLESEHCAGTAGIISRKAMFFLDRAVLPPELHPVDLMPGNP 189
 ID 09CLR8 PRELIMINARY; PRT; 208 AA.
 AC 09CLR8;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Hypothetical protein Pm141.
 GN Pm141.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Pasteurella.
 OX NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PM70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida Pm70.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 DR EMBL; AE006155; AAK03225.1;
 DR InterPro; IPR002654; GT_25.
 DR Pfam; PF01755; Glyco_transf_25; 1.
 DR Hypothetical protein; Complete proteome.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 208 AA; 24649 MW; 29625BA115E4047A CRC64;
 Query Match 23.9%; Score 349; DB 16; Length 208;
 Best Local Similarity 37.3%; Pred. No. 4.7e-24;
 Matches 82; Conservative 40; Mismatches 74; Indels 24; Gaps 6;

OY 61 ALPNVVS-YCGRDFLALDENHGTAGIISGAAYLLEIFKNNMNNIPIIDHILIFRF 119
 ID 09CLR8 PRELIMINARY; PRT; 208 AA.
 AC 09CLR8;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Hypothetical protein Pm141.
 GN Pm141.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Pasteurella.
 OX NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PM70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida Pm70.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 DR EMBL; AE006155; AAK03225.1;
 DR InterPro; IPR002654; GT_25.
 DR Pfam; PF01755; Glyco_transf_25; 1.
 DR Hypothetical protein; Complete proteome.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 208 AA; 24649 MW; 29625BA115E4047A CRC64;
 Query Match 23.9%; Score 349; DB 16; Length 208;
 Best Local Similarity 37.3%; Pred. No. 4.7e-24;
 Matches 82; Conservative 40; Mismatches 74; Indels 24; Gaps 6;

OY 120 LAGEELAWYQSLSPALCICIEL--QLNENSLDSQLESEKRYRLAERAKRKTKYREKY 176
 ID 09CLR8 PRELIMINARY; PRT; 208 AA.
 AC 09CLR8;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Hypothetical protein Pm141.
 GN Pm141.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Pasteurella.
 OX NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PM70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida Pm70.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 DR EMBL; AE006155; AAK03225.1;
 DR InterPro; IPR002654; GT_25.
 DR Pfam; PF01755; Glyco_transf_25; 1.
 DR Hypothetical protein; Complete proteome.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 208 AA; 24649 MW; 29625BA115E4047A CRC64;
 Query Match 23.9%; Score 349; DB 16; Length 208;
 Best Local Similarity 37.3%; Pred. No. 4.7e-24;
 Matches 82; Conservative 40; Mismatches 74; Indels 24; Gaps 6;

OY 243 NTFKRLRLA--LTGIREREKRORREQLIGKIIVPQ 279
 ID 09CLR8 PRELIMINARY; PRT; 208 AA.
 AC 09CLR8;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Hypothetical protein Pm141.
 GN Pm141.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Pasteurella.
 OX NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PM70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida Pm70.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 DR EMBL; AE006155; AAK03225.1;
 DR InterPro; IPR002654; GT_25.
 DR Pfam; PF01755; Glyco_transf_25; 1.
 DR Hypothetical protein; Complete proteome.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 208 AA; 24649 MW; 29625BA115E4047A CRC64;
 Query Match 23.9%; Score 349; DB 16; Length 208;
 Best Local Similarity 37.3%; Pred. No. 4.7e-24;
 Matches 82; Conservative 40; Mismatches 74; Indels 24; Gaps 6;

Wed May 28 09:28:19 2003

Search completed: May 27, 2003, 18:17:18
Job time : 90 secs

us-10-007-267a-8.rspt

Page 7

THIS PAGE BLANK (USPTO)